Essays in Experimental Economics

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

Katharina Sabine Momsen

Sommersemester 2017

Abteilungssprecher: Prof. Dr. Jochen Streb
Referent: Prof. Dr. Henrik Orzen
Korreferent: Prof. Stefan Penczynski, Ph.D.

Tag der mündlichen Prüfung: 15. August 2017

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig angefertigt und die benutzten Hilfsmittel vollständig und deutlich angegeben habe.

Mannheim, 10. Juli 2017 Katharina Momsen

Acknowledgments

First and foremost, I would like to thank my supervisor Henrik Orzen for his continuous support during my PhD studies in Mannheim. I am very grateful for the productive discussions about my research, the constructive criticism and the highly valuable advice on the chapters in this dissertation. I have learnt a lot from him, especially while working on our joint project. I am also very grateful to my second supervisor Stefan Penczynski for his support, the highly useful comments on my research and the helpful advice on economic and career related questions.

Furthermore, I want to thank the members of the HeiKaMaX community and the experimental economists at ZEW for giving me the opportunity to present my research. Thank you for your valuable comments. The regular workshops and seminars have strengthened my interest in behavioral and experimental economics.

A special thanks goes to the experimental economists at the University of Heidelberg, especially Jörg Oechssler, for giving me the opportunity to use their lab for my sessions. I also want to thank Andrej Svorenčík for managing the mLab in Mannheim and helping me with the organization of my sessions.

I would like to thank my fellow doctoral students in Mannheim for the enjoyable time we had together. In particular, I want to mention the best office mates one can imagine, Florian and Jasper.

I also want to mention the members of my rowing team. After (or before) a long day in the office it was great to spend time with you on the Neckar.

Last but not least, I wish to thank Johannes, my brothers Björn and Felix and my parents for always encouraging me and supporting me. Thank you.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

Consumers make a multiplicity of purchase decisions in their daily lives. However, oftentimes they are unfamiliar with the specific characteristics of goods and services or the market conditions in general. Hence, they regularly end up making suboptimal choices for a wide variety of reasons: they may lack time, motivation or mental capacities to familiarize themselves with each decision. As a consequence, they may suffer welfare losses from suboptimal choices. In particular, they may fail to identify and select their preferred variant when several versions of a product exist. Furthermore, due to limited price information they might purchase overpriced goods which would have been available at a cheaper price. With little overview of the market they might not be able to purchase at all, because they decide to visit a store where the product is out of stock. On top of that, consumers' purchasing behavior may exhibit boundedly rational patterns: they may fail to draw the correct consequences from their information or fail to correctly interpret their experience with some products. Research on actual decision making in purchase situations is especially important since firms may exploit consumers' behavior for their own benefit which often leads to welfare losses on the consumer side.

It is difficult to identify all the above mentioned phenomena in the real world. Economic experiments are a useful tool to contribute to a better understanding of decision making in an industrial organization context: they permit empirical analyses when real world data are scarce or of poor quality. Furthermore, they allow the researcher to completely control both the setting and the potential influencing factors. The experimental design may even allow her to observe the counterfactuals of the decisions – a feature that data from outside the laboratory often times do not offer. Endogeneity issues that may severely bias findings derived from real-world data are absent from experimental data as treatments are imposed by the researcher. In addition to observing the final decisions, experimental economists can adjust the design of their experiments such that the decision making process becomes – at least partially – tractable. Therefore, I am convinced that experiments in industrial organization contribute to a better comprehension of the functioning of consumer markets.

This thesis contains three self-contained articles, each studying decisions made under incomplete information among buyers with the help of experiments. While Chapter 2 and Chapter 3 study the interaction of buyers and sellers in different market settings, Chapter 4, which is joint work with Henrik Orzen, focuses on consumer decision making. Chapter 2 analyzes markets for horizontally differentiated credence goods in which competing sellers have the option to give buyers non-verifiable recommendations. In Chapter 3 I investigate the influence of better price information and the possibility to communicate among buyers on prices and welfare in posted offer markets with capacity constrained sellers. Chapter 4 seeks to identify decision making heuristics used by buyers in an environment with very limited information both from a one-shot perspective and over time.

In Chapter 2 I study a credence goods market with horizontal product differentiation. Consumers have different valuations for the available product versions as it is often the case, e.g. in today's consumer electronics markets, but these valuations are unknown to them. If sellers decide to invest, they can observe these valuations. Yet investment is costly and cannot be observed by buyers. In order to influence the buyers' purchase decisions, the seller can recommend one of the versions. However, buyers do not know if the seller has invested and is thus capable of giving valuable advice. In addition, they cannot observe if the advice aims at maximizing their payoff or the seller's profit. In this setting I want to find out if sellers decide to invest to give expert advice and if the recommendations can be trusted by consumers. Furthermore, I ask if two policy measures – the introduction of watchdogs and the revelation of the sellers' investment decisions – influence the quality of the recommendations. In order to answer these questions I design an experiment in which subjects interact repeatedly on markets taking the roles of buyers and sellers. In the treatments with watchdogs, the sellers' recommendations are judged from time to time and the results are presented to buyers. If the investment decision is revealed, it is presented to buyers when they need to choose which seller they want to interact with. I find that in the absence of watchdogs and with concealed investment decisions, investment rates are low. Even if sellers have invested, they give selfish advice. However, the low price level ensures a relatively high consumer surplus. Revealing the investment decisions increases the share of investing sellers, yet it does not increase the quality of the recommendations. Only with both watchdogs and revealed investment decisions is the quality of advice increased and buyers accept higher prices.

Chapter 3 deals with the effects of communication and price information in posted offer markets. In the considered market constellation, capacity constrained sellers face more buyers than they can jointly serve. Thus, miscoordination among buyers can arise and too few transactions might take place. I investigate how a larger share of buyers with price information influences the prices. In addition, I want to find out if it is beneficial for buyers to be able to communicate their purchase intentions before they decide with which seller they want to interact. Increasing the share of informed buyers might lead to ambiguous effects of information on prices: on the one hand, when more buyers are price-sensitive each seller might want to decrease his price because he can rely less on random purchases by buyers who do not observe the prices. On the other hand, consumers know that the other informed consumers will also try to purchase from sellers with relatively low prices. Thus, competition for the chance to purchase from a cheap seller increases when more consumers are informed about the prices. Consequently the expected payoff of purchasing from a cheaper seller decreases as the probability of obtaining the good is lower. In the experiment I seek to find out which effect prevails. Furthermore, I investigate the effects of buyer communication: communication might increase consumer surplus because buyers can better coordinate their purchases. Yet the achieved coordination may also have a negative effect on buyer surplus if sellers raise their prices trusting that they will always be able to sell their good. I investigate these effects in a market experiment with repeated interaction. I observe lower prices in the treatment with a larger share of informed buyers when buyers cannot communicate. Surprisingly, communication alters the effect of information. It induces higher levels of coordination among buyers which provokes sellers to raise their prices. Overall, buyers suffer from the possibility to exchange their intentions before making a purchase decision.

In contrast to the other two chapters, Chapter 4, which is joint work with Henrik Orzen, considers only the demand side. It is inspired by Spiegler's (2006) theoretical work on the market for quacks. He considers a situation in which consumers need a certain product – in his context remedies against an illness –, but they are unfamiliar with the market and the available products. According to Spiegler, the natural approach to reach a decision in this setting is to gather anecdotes about the available products and about the consequences of not purchasing. He assumes that consumers purchase the cheapest product which has received a positive anecdote. Based on this assumption Spiegler shows that markets for useless products – products that are equally effective as the option of not purchasing – can emerge. Crucially, his results depend on actual consumer behavior in these types of situations. With the help of an experiment we want to shed light on the actual purchasing behavior of poorly informed consumers. We develop a real effort task with the intention to create a situation resembling Spiegler's context. We implement two versions of the real effort task: an easy version and a difficult version. When subjects are confronted with the difficult version, they can purchase a costly product which might transform the difficult version into the easy one. However, it may also become simple without purchasing. In a between-subject design we alter the effectiveness of the costly products and of the option not to purchase. Subjects play the real effort task repeatedly, allowing us to observe how their behavior changes when they become more experienced with the situation in general. Subjects are offered different products in each round such that their experience cannot be related to specific products. The only information about the products they receive are prices and anecdotes. A positive anecdote informs them that the product would have been successful in simplifying the difficult real effort task in a previous round. We find that subjects are very price-sensitive and lose some of their trust in anecdotes as they grow more experienced. There is very little learning related to the effectiveness of the products relative to the option not to purchase. Even when the costly products are superior to the option not to purchase, a substantial share of the experienced subjects does not seek to purchase. At the same time, a large share of subjects purchases products which are as effective as the option not to purchase. Overall, our results suggest that some demand for ineffective products persists when consumers lack information about the market.

Chapter 2

Recommendations in Horizontally Differentiated Credence Goods Markets*

2.1 Introduction

Consumers need to make a plethora of purchase decisions in their daily lives. Due to the large amount and the diversity of decisions, they may lack time, interest, motivation or mental capacities to reflect on each purchase decision very deeply and familiarize themselves with each product they need. Especially when several different variants of each product exist, consumers may not know which version to choose. Even though they serve the same general purpose, versions may induce different utility levels for each consumer because of different personal preferences. When deciding which version to purchase, a poorly informed consumer may find it very difficult to identify the one that maximizes her utility if she is not familiar with the specific characteristics.

Complex products such as financial services[1], consumer electronics or cars may be hard to assess for some consumers. Take for example electronic devices like smartphones: instead of going through the troublesome process of informing themselves about all potential options of the product they need, consumers often rely on the advice from retailers. Yet the quality of the advice they deliver may differ: those who employ better

[1] In a recent study, the German consumer organization "Stiftung Warentest" analyzed the investment advisory services of 23 financial institutes and discovered that only three provide good advice (see https://www.test.de/Anlageberatung-Nur-3-von-23-Banken-beraten-gut-4964413-0/).

* I am grateful to the seminar audience at the HeiKaMaxY in Karlsruhe for their valuable comments and suggestions.

trained or more experienced salespeople can better assess the consumers' utility from each variant and can thus give better recommendations. As better trained salespeople require higher wages, their employment constitutes an investment. Observing the recommendation, it may be hard for consumers to judge if it stems from a qualified salesperson and whether the advice is genuine or whether it serves to maximize the seller's surplus. After the purchase the consumer will experience her consumption utility, but she may never find out if her choice was optimal because her valuations for the alternative options remain unknown to her. This lack of information may be exploited by sellers who can gain by giving selfish recommendations. Hence, there are situations in which consumer electronics fall into the category of "credence goods" (Darby and Karni, 1973): before purchasing, consumers do not know which product to choose and in contrast to "experience goods" they still lack information about the relative goodness of their choice after the purchase.

So far, both theoretical and experimental papers have mainly investigated credence goods from a vertical product differentiation perspective with two versions of a good – a small version that satisfies the consumer's need with some probability and a large version which is always sufficient for the consumer. The consumer does not know which variant she needs – unlike the supplier – and after the purchase she does not find out which version she has received. This information asymmetry can be exploited by the seller who can either provide the small version when the larger one is needed (undertreatment), provide the larger version although the smaller one would have been sufficient (overtreatment) or provide the small version, but charge for the large version (overcharging). Although the existing theoretical, empirical and experimental literature (see the next section for an overview) is suitable to describe many credence goods problems very well, markets with a *horizontal* product differentiation component such as consumer electronics have not been investigated yet. In these markets *mistreatment* instead of *over-* or *undertreatment* may reduce the consumers' welfare and in contrast to undertreatment sellers can expect their mistreatment to remain undetected, because consumers will usually not find out if another version would have been better for them. In this paper, I analyze credence goods markets with horizontal product differentiation using a laboratory experiment.

The experimental setting resembles markets in which consumers repeatedly interact with sellers by purchasing different products. Consider again the consumer electronics market: in many countries there are only a few different retailers which all offer the same variants of the same products. Thus, the retailers only differ in their service quality and in their prices. Each time you need a new electronic device, you need to choose among the same retailers. One day, you may need a new smartphone. On another day, your fridge might be broken and you are looking for a replacement. When deciding which retailer to visit, you recall your experience from the past purchase: did the retailer recommend a product variant that made you content? Or did he suggest a version that turned out not to match your needs satisfactorily well? Hence, the past recommendations may influence today's purchases.

In the experiment subjects interact repeatedly in markets consisting of eight subjects, four taking the role of buyers and four acting as sellers. Markets persist in the same composition throughout the whole experiment. All sellers in a market offer the same range of product variants with each seller facing different costs for each variant in order to simulate the realistic scenario that different variants create different profit margins. Consumers differ in their valuations for the variants. The valuations are unknown to them such that they cannot identify their most preferred variant. Sellers can help buyers to select a variant by giving them a recommendation. This recommendation is not necessarily consumer-friendly, nor is it always based on the consumer's preferences. Without investing into trained personnel, sellers are unable to observe the consumers' valuations for the different product versions; they only observe their own costs. Only if they invest, can they observe their clients' valuations for each variant together with their own costs. Investment is costly and must be renewed in each round. In the baseline treatment consumers do not know which sellers have invested. As each subject is informed about her earnings, buyers can infer their valuation from the purchased variant, but they do not find out if they would have had higher valuations for other variants such that they cannot judge the quality of the recommendation. If sellers do not invest, consumers cannot rely on the recommendations and may suffer from purchasing suboptimal versions. I seek to identify factors which potentially influence the sellers' investment behavior and the quality of the recommendations. As one treatment variation I introduce watchdogs who test and rate the quality of the sellers' advice. As another treatment variation I reveal the sellers' investment decision to consumers.

The potential presence of watchdogs resembles occasional checks by consumer organizations such as the "Which?" Consumers' Association in the UK, Consumer Reports in the USA and Stiftung Warentest in Germany. These organizations test products and services thoroughly to better inform consumers about product attributes and detect fraud towards consumers. In addition, tests conducted by magazines and online platforms complement those performed by the established watchdogs. The test results are revealed to consumers and help them choose products and services based on features they would normally not have been informed about. Sellers know that these organizations exist, but they do not know if and when their products or services will be subject to a test. It is only revealed ex post if they have been tested and how their product or service has performed. Similarly, in the experiment both buyers and sellers know that the recommendations are subject to occasional checks in randomly determined rounds, but they cannot identify these rounds ex ante.

A visible investment decision resembles markets in which consumers can observe whether sellers have undertaken any training or have earned specific qualifications in order to be able to fulfill their jobs properly. This treatment variation resembles objective proofs of the sellers' qualification such as certificates placed prominently in stores or offices. Even though consumers know that the seller is capable of providing the appropriate service, they do not know if he actually makes use of his acquired skills to increase their consumer surplus. Instead, sellers might also act selfishly and recommend their most profitable variant not taking into account the information they have about the consumers' preferences.

Using an experiment I seek to find out how often sellers invest in giving valuable recommendations, what prices they set and according to which heuristic they select the variant they recommend: do they aim at maximizing their own surplus, the consumer's surplus or total welfare? Concerning consumers' behavior I ask if higher prices are interpreted as a signal that the seller has invested or whether consumers seek cheap sellers. I also want to find out how buyers react to recommendations – do they follow them or do they rather ignore them? I investigate whether the revelation of the investment decision has an impact on the market outcome. Furthermore, I ask if the occasional presence of watchdogs improves consumer welfare.

I find that in the absence of watchdogs and without observable investment decisions few sellers invest and most recommendations are selfish. Prices are cheap and consumers do not learn anything about their most preferred version from the recommendations. In spite of this, they mostly follow the recommendations. The low price level ensures relatively high payoffs for consumers even though they usually do not purchase their most preferred option. The visibility of the investment decision encourages more sellers to invest, but does not reduce the selfishness of the recommendations. Potential visits by watchdogs increase the consumer-friendliness of the recommendations, but only the joint presence of watchdogs and the revelation of the investment decision provoke frequent investment *and* consumer-friendly advice. In general, consumers are very price-sensitive purchasing mostly from the cheapest seller, yet in the treatment with watchdogs and observable investment decisions they accept higher prices. Thus, not only buyers do benefit from better recommendations in this treatment, but also sellers can charge higher prices and therefore recover their investment costs.

The remainder of this paper is organized as follows: in section 2, I will summarize the related theoretical and experimental literature. Then, I will explain my experimental design and in section 4, I derive my predictions and main research questions. I will present my results in section 5. The last section concludes and the appendix provides a translation of the instructions.

2.2 Related Literature

The seminal paper on credence goods was written by Darby and Karni (1973). They extended the categories of goods by adding credence goods to ordinary goods, search goods and experience goods. Credence goods markets are characterized by large information asymmetries between sellers and buyers: while the seller is an expert about the good (or the quality of the good) the consumer needs, the consumer lacks this information. In contrast to experience goods, the consumer cannot fully assess the trade even after purchasing. As typical examples for credence goods Darby and Karni (1973) mention taxicab rides in unknown cities and medical treatments. They argue that fraud at the expense of consumers occurs in credence goods markets due to the fact that sellers both perform the diagnosis and choose the treatment. Sellers can undertreat, overtreat and overcharge the consumer – depending on the market environment and the financial incentives. In my setting, experts also perform the diagnosis, but instead of immediately providing the treatment, they rather give a recommendation which buyers can choose to follow. Furthermore, with horizontal product differentiation consumers suffer from mistreatment instead of over- or undertreatment.

Pesendorfer and Wolinsky (2003) introduce the possibility to gather multiple opinions to credence goods markets. In their model consumers can choose between competing experts whose diagnostic effort is unobservable and costly. Undertreatment is possible as the success of the treatment is not contractable. They find that price competition reduces the experts' incentive to exert effort and conclude that interventions to limit price competition may increase welfare in credence goods markets. I also focus on a setting with harsh price competition and analyze factors that influence the sellers' willingness to invest, but I exclude the possibility of consulting several sellers.

In a theoretical paper Dulleck and Kerschbamer (2006) provide an extensive overview of the literature on credence goods. Developing a general model they seek to unify the existing theoretical literature. Similar to Emons (1997) they find that the price mechanism can ensure nonfraudulent behavior when certain conditions are met: to ensure the existence of equilibria without fraud, consumers need to be homogeneous – which is not fulfilled in my setting –, economies of scope between diagnosis and treatment have to exist and either the type of the provided treatment has to be verifiable or the seller needs to be liable for the outcome.

Dulleck and Kerschbamer (2009) analyze competition between experts and discounters in credence goods markets. In contrast to discounters, experts have invested in diagnosis effort to identify the consumer's problem. Similar to the baseline treatment of my experiment, investments are costly and unobservable. Consumers can free ride on an expert's diagnosis, i.e. they can ask the expert for advice, but purchase from a cheaper discounter. In addition to this free-riding problem on the consumer side, there exists a moral hazard problem for experts: they do not necessarily provide an honest diagnosis. In the model, the final success of the service is observable and verifiable and only experts provide insurance against failure. When switching costs are sufficiently low, there is no equilibrium in which experts always perform a diagnosis before recommending one version, instead they randomize between always providing the cheap version and giving honest recommendations. This strategy ensures that consumers are not completely certain about their actual needs and thus prefer experts who provide insurance against insufficient treatment. While my experiment also features the potential coexistence of expert sellers and discounters in a market, I do not allow for second opinions and thus exclude the free-riding problem. Furthermore, my experts do not insure against mistreatment.

Turning to the experimental literature on credence goods, the large-scale experiment by Dulleck et al. (2011) studies the determinants of efficiency in credence goods markets and thereby tests the theoretical predictions of Dulleck and Kerschbamer (2006). Whereas in theory either liability or verifiability ensures efficiency, they find that in the experiment verifiability has only a minor effect, but liability a major impact on efficiency. Furthermore, they find almost no effect of reputation on the market outcome and observe significantly lower prices when sellers compete for consumers instead of being matched to certain consumers. Yet competition does not yield efficiency.

In another experimental paper, Dulleck et al. (2012) examine the relationship between prices and quality. They ask whether good experts who provide sufficient treatment set high prices or whether high prices induce sellers to provide sufficient quality. They find that good experts make use of high prices to signal their quality, but find no evidence for the opposite reasoning. Bad sellers rather set high prices to mimic good sellers, but deliver unsatisfactory quality. In my experiment, sellers may also use prices to signal their investment decision.

Mimra et al. (2016a) analyze how price competition and consumer information about past expert behavior influence fraud in a setting where undertreatment and overcharging is possible. They observe more fraudulent behavior under price competition than under fixed prices, concluding that the price decline under price competition inhibits quality competition as the price pressure undermines reputation building. Whether consumers only observe the private history of their own decisions and the respective outcomes or if consumers are also informed about the histories of other consumers in their market does not have an impact on fraud.

A second paper by the same authors (Mimra et al., 2016b) examines the possibility to gather second opinions in a market for expert services experimentally. Although second opinions are per se inefficient because diagnosis costs and search costs occur twice, they can be beneficial in markets where sellers diagnose and provide services, because they may reduce the level of overtreatment. The authors find that lowering the search costs induces buyers to seek more second opinions such that overtreatment decreases. Thus, market efficiency rises under the introduction of second opinions, because the reduction in treatment costs exceeds the increase in search costs.

Schneider (2012) examines credence goods in a field experiment. He takes a deliberately damaged car to several mechanics either stating that he plans to move away or leaving a home address close to the garage. In the former case, the mechanic should have less reputational concerns than in the latter. However, the author does not find an effect of reputation on the level of fraud. Balafoutas et al. (2013) examine fraud related to taxicab rides in Athens. They observe that passengers who appear to have little information about the shortest route are taken on longer detours. Passengers who are unfamiliar with the local tariff system are more likely to receive a manipulated bill.

Another strand of the literature related to credence goods is the experimental literature on trust games. The difference between trust games and credence goods lies in the information asymmetry between buyers and sellers. Trust games can be interpreted as markets for experience goods. Inefficiencies related to trust games can result from undertreatment or no market interaction, yet in credence goods markets consumers may additionally suffer from overtreatment and overcharging such that the outcome of a transaction is a weaker signal for the quality of the treatment. Huck et al. (2012) find that enabling sellers to build a reputation increases trust, but giving trustors the possibility to choose between trustees has a larger positive impact on trust. When trustees can determine the payoffs of rewarding or exploiting trust, trust and trustworthiness are lower compared to a setting with pre-determined payoffs (Huck et al., 2016). Hence, price regulation as opposed to price competition may increase welfare in the context of experience goods.

Finally, the current paper relates to the literature on ultimatum games. In ultimatum games the proposer suggests how to split a pie and the responder can accept the split, in which case both players receive the suggested shares or reject the trade such that both players receive nothing. A rational responder should never reject as she benefits from accepting any non-negative offer. In my setting, the seller announces a price and thus suggests how to split the pie which consists of the consumer's valuation minus the seller's costs. The standard form of the ultimatum game has been analyzed in the laboratory thoroughly (see e.g. Güth et al., 1982), but also several extensions of the standard game have been taken to the lab. Mitzkewitz and Nagel (1993) study a version in which the size of the pie is only known to the proposer. It can lie between one and six and is determined randomly. They alter whether the responder is informed about the proposer's offer for the responder or about the amount the proposer claims from himself. In the former case, the proposers deciding about the split of a large pie try to pool with proposers with smaller pies offering less than half of the pie. In the latter case, proposers demand half of the pie when it is large. They demand three when the cake is smaller and they demand the whole cake when it is smaller than three. In my setting, buyers only observe the price sellers demand. As they do not know the actual realizations of costs and valuations, they can only vaguely assess the split imposed by the price.

Another version of the ultimatum game is the yes-no-game by Gehrig et al. (2007). In this game the responder does not observe the proposal, but in contrast to the dictator game she can still accept or reject the unseen offer. The authors find that the offers resemble offers in dictator game and are significantly lower than offers in ultimatum games.

In Kriss et al. (2013)'s version of the dictator game, it is again only the proposer who can observe the size of the pie. They alter the possibilities of deception: when implicit deception is possible, sellers can make offers that would be fair if the pie size was different. Explicit deception allowed lies about the actual pie size. They find offers to misrepresent the actual cake size more when explicit deception was possible. The deception in my experiment relates to the recommendations sellers give. They can claim that one variant benefits the buyer most when it is in fact the version that maximizes their profits.

A closely related paper by Anbarcı et al. (2015) examines ultimatum games with messages and offers. While messages are cheap talk and can be false, offers state the actual split of the pie. As opposed to messages which always arrive at the responder, offers are only observed with some probability. When this probability increases, offers increase, messages overstate the offers less and responders are more likely to accept even if they do not observe the offer. The authors find that the likelihood with which a deal is accepted increases when responders notice that the message has been honest and when messages have been sent.

Corazzini et al. (2014) extend the ultimatum game with messages and offers by introducing several competing proposers. They suggest political campaigns by politicians who do not necessarily keep their promises as example. The messages sent by politicians resemble the sellers' recommendations in my experiment. However, a recommendation cannot be identified as a lie because buyers do not know if other versions would have yielded higher levels of utility which makes the problem more realistic, but less tractable. Similar to Corrazzini et al. (2014) my setting includes competition among proposers, yet their proposers compete in messages while mine compete in prices and only later send messages. This difference is due to the different applications of the experiment, i.e. political campaigns vs. product recommendations.

2.3 Experimental Setup

The experiment consists of four between-subject treatments implemented in a 2x2 design: I alter both the visibility of the investment decision and the presence of watchdogs (see Table 2.1). In the baseline treatment, no watchdogs are present and the investment decision remains unobservable to buyers. In the 'W' treatment watchdogs appear in randomly selected rounds to openly judge the quality of sellers' recommendations and in the 'O' treatment buyers observe the investment decisions before they select a seller. In the forth treatment, 'O+W', both watchdogs can occur and the sellers' investment decisions are revealed. The experiment was programmed in zTree (Fischbacher, 2007) and it was run at the University of Mannheim in the mLab in Spring 2017. Subjects were recruited using ORSEE (Greiner, 2004). In total 175 subjects participated and earned on average 13.70 €. Most subjects were undergraduate students at the University of Mannheim from all fields. The sessions lasted on average 80 minutes and consisted of 30 rounds.

Table 2.1: Treatments

		Observable investment decisions?	
		No	**Yes**
Watchdogs present?	**No**	Baseline	Observability
	Yes	Watchdogs	O + W

Subjects are split into groups of eight subjects under a fixed matching protocol. A group represents a market consisting of four sellers and four buyers. Subjects are randomly assigned to their roles at the beginning of a session and keep their role. While buyers cannot be identified by sellers, each seller is always represented by the same number, e.g. the same participant is always referred to as 'seller 1' throughout the whole experiment. This feature allows sellers to build a reputation based on each buyer's private history. In the existing experiments on credence goods, sellers usually offer two versions of a good or service: the large and possibly more expensive version always satisfies the buyers' needs, whereas the cheaper version is only sufficient with a certain probability. In order to introduce horizontal product differentiation to a credence goods setting, I increase the number of product versions and ensure that sellers' marginal costs and buyers' valuations are uncorrelated. Each seller offers the same variants, but the variants change from round to round to ensure that learning corresponds to the situation in general, not to specific products. Buyers differ in their valuations for the different versions. As they learn about the net utilities they derived from their purchases, but remain uninformed about those of the other variants, they can build only a vague opinion about the quality of their seller's recommendation. Furthermore, they remain unfamiliar with the products and cannot transfer any knowledge about the products from round to round.

In the experiment sellers are not capacity constrained. Each variant has different marginal costs which occur only when it is sold. These marginal costs are integers randomly drawn for each seller without replacement from a discrete uniform distribution between 0 and 10 Experimental Currency Units (ECUs) and differ among sellers. In reality, these costs might depend on the sellers' contracts with the producers and differ depending on the negotiated terms. Buyers have different valuations for the five product variants. A buyer's valuations are integers drawn without replacement from a discrete uniform distribution between 10 and 20 ECUs.[2]

[2] Based on the integral from which the valuations are drawn, one can derive the expected value of the largest integer in a draw: if five integers k were picked without replacement from an interval $\{1, ..., n\}$ and I is the value of the highest integer, then $E[I] = \frac{k}{k+1}(n+1)$. For the buyers' valuations the interval $\{1, ..., 11\}$ needs to be shifted to the left by nine integers, as the lowest possible valuation is 10. Thus, $E[I] = \frac{5}{6} * 12 + 9 = 19$, i.e. in expectation, buyers have a valuation of 19 for their most preferred version

If a seller decides to invest, he incurs additional costs of 1.5 ECUs which need to be paid even if he cannot sell anything in the current round. I chose 1.5 ECUs as investment costs, because this cost level renders the investment decision welfare neutral: by investing, sellers learn their buyers' valuations and can thus include this knowledge in their recommendations which increases expected welfare by 1.5 ECUs.

In the treatments with watchdogs both buyers and sellers know about the occasional visits by watchdogs, but they cannot identify the relevant rounds ex ante. Note that in the experiment it is not the products themselves, but the *quality of the recommendations* which is tested. As all sellers offer the same product variants, each seller could try to attract buyers by giving better recommendations. Yet buyers cannot always detect if the recommendation has been optimal for them. Only watchdogs can find out if the seller has recommended the optimal product variant. I implement watchdogs by randomly selecting four test rounds from the 30 rounds of the experiment. However, I ensure that the last five and the first three rounds are excluded: I rule out the very first rounds as I wanted subjects to gain some experience before being confronted with watchdogs. The last rounds are not included either because watchdogs were designed to facilitate reputation building which would be pointless at the end of a session.

In a test round, one recommendation from each seller who was successful in attracting a positive number of buyers is selected randomly and graded. The grading scheme is designed as follows: When proposing the variant that yields the lowest consumer surplus, the seller receives the worst possible grade (5). Analogously, when the seller recommends the variant which maximizes the consumer's utility, his grade is a '1'. In case he has given a recommendation which neither minimized nor maximized consumer surplus, the grade is a '3'.[3] A seller is not graded (0) if he has not found a consumer or has opted not to give a recommendation. If a seller has found more than one consumer and has given one consumer friendly recommendation and one recommendation that minimizes the other consumer's utility, only the randomly determined recommendation is taken into account and the other does not influence his grade. In the subsequent round, consumers are informed about the grades of all four sellers in their market before they select one of the sellers. A consumer is never informed whether the sellers' grade is based on her own previous transaction with the seller. She can only infer that it must have been her interaction when all four sellers are graded, because then all four sellers have found

in each round. If they derive an average valuation lower than 19 from their payoffs, they might expect that they have not purchased their preferred versions in the past rounds.

[3] The grading system is inspired by German school grades. However, I exclude the grades '2' (good) and '4' (sufficient) in the experiment.

exactly one consumer each.

The investment decision is revealed when buyers need to choose between the sellers in their market. In the treatments with hidden investment decisions, they only observe the sellers' prices, so the investment decision is added as an additional piece of information in the other treatments.

At the beginning of a round, sellers decide whether they want to invest in order to gain the ability to observe their customers' valuations. Then, they set one price for all five product variants.[4,5] The marginal costs of each product variant are depicted on their decision screens. As soon as all sellers in a market have determined their prices, buyers observe them and individually select the seller they want to interact with. Depending on the treatment, buyers do or do not observe which sellers have invested. Furthermore, they are shown their past purchase decisions and the respective decision outcomes. In the treatments with watchdogs, the sellers' grades from the past test rounds are depicted. When all buyers have determined their interaction partner, sellers are informed which buyers they were able to attract. For each variant, they are reminded of their margin (price − marginal costs) and, if they have invested, they can observe their clients' net utility (valuation − price). Based on these pieces of information, they decide for each of their customers individually which variant to recommend or if they prefer not to give a recommendation. Next, buyers receive the advice and, in the treatments with observable investment decisions, they are reminded whether their seller has invested. Buyers can follow the recommendation, they can decide to purchase a different variant or they can opt out and decide not to purchase at all in this round. At the end of a round, all subjects are informed about their earnings. For a seller, the earnings in a round are determined by the price they set from which the marginal costs of the variant they sold are subtracted. If they found more than one customer, their earnings from the transactions are added. Investment costs are subtracted only once. A consumer earns her net utility (valuation - price) if she decided to purchase and zero if she has not interacted with a seller. Then a new round begins in which buyers receive new identification numbers and new products are traded, i.e. new costs and valuations are randomly determined.

[4] For simplicity, sellers do not set a single price for each variant. The price they set is supposed to resemble the average price level of the seller.

[5] Sellers had to set their price within 45 seconds. When they forgot to set their price within the time limit and ignored the reminder, the computer automatically saved a price of 20 ECUs such that the experiment could continue. I exclude the very few rounds in which subjects failed to set a price from the analysis.

A the end of the experiment, ten of the 30 rounds are selected randomly for payment. The average earnings of these ten rounds are converted into Euros using different exchange rates for buyers and sellers to ensure comparable earnings. For sellers, one ECU was multiplied by five to get a €1, whereas for buyers the exchange rate was one. Additionally, buyers received a fixed fee of €2.50 and sellers of €5. On average, sellers earned €12 and buyers €14.40.

2.4 Predictions and Research Questions

I want to understand markets for credence goods with a horizontal product differentiation component. I therefore ask whether buyers follow the recommendations of the sellers and I differentiate between situations in which buyers know about their sellers' investment decisions and situations in which they remain uninformed. Besides, I seek to find out according to what objectives sellers decide to give a recommendation and which variant they select. Do they only give recommendations when they have invested? Do they recommend the variant that maximizes their profit or do they also take into account the buyer's welfare? Does their decision depend on the presence of watchdogs? Furthermore, I analyze how buyers select a seller. Do they select the cheapest seller? Are they loyal to their seller if they earned a high payoff and do they switch if their payoff in the previous round was low? Do they make use of additional information if available, i.e. do they take the sellers' grade and investment decisions into account when determining their interaction partner? Here, I am curious to find out how sellers react to the buyers' purchase behavior, what prices they set and if they decide to invest.

Since each market consists of four sellers, competition to attract consumers is fierce. Sellers compete in two dimensions: prices and the quality of recommendations. Thus, they can increase their attractiveness for consumers both by undercutting their competitors' prices and by giving better recommendations. However, to be able to give good recommendations they need to invest, which is costly. When sellers are selfish, they would only be willing to invest if this strategy paid off for them, i.e. if it justified higher prices or helped them build a good reputation to attract more consumers in future rounds.

In the baseline treatment, buyers cannot observe the sellers' investment decisions. If sellers try to signal their investment through a higher price, other sellers can easily mimic their behavior by also charging higher prices without having spent any money on investing. Therefore, it should be impossible to benefit from a positive investment decision by charging a higher mark-up and passing on the investment costs to consumers. Investing with the intention of giving good recommendations and thereby convincing buyers to come back repeatedly is also a rather unlikely strategy in this setting: as buyers can only infer their valuation for the purchased variant, they often do not find out if they actually had a higher valuation for another variant.[6] As consumers lack part of the information necessary to entirely judge the received recommendation, they might not always reward the seller for a consumer-friendly recommendation by purchasing from him again in the subsequent round. Hence, sellers are likely to find it difficult to build a reputation of giving good advice which decreases their incentive to invest. Anticipating that sellers have most likely not invested, consumers focus on prices. Given that all sellers offer the same variants, consumers should choose to interact with the cheapest seller.

Under the assumption that each seller wants to maximize his profits he tries to attract all consumers in his market by slightly undercutting his rivals' prices. As he cannot be sure which version his client will purchase and therefore assumes that she chooses randomly, his average costs matter for his price setting decision. Knowing the distribution from which costs are drawn, the seller expects his rivals' average costs to be equal to 5. The seller knows that the other sellers will set prices of 5 ECUs on average if they want to break even. Thus, he anticipates to find no consumers at prices exceeding 5 ECUs, to attract all consumers with a weakly positive probability when he charges a lower price and to find on average one consumer when he charges a price of 5. He might take into account his own costs in the current round. In rounds with very low costs, he might decide to undercut the price of 5 to attract more consumers, with moderate costs he sets a price of 5 and with high costs, he sets a price above 5 ECUs expecting profits of zero.

In the treatments with observable investment decisions, consumers know which sellers are *capable* of giving buyer-friendly recommendations. Yet, they do not know if sellers actually make use of their knowledge about buyers' valuations. The above mentioned difficulty for sellers to build a reputation remains unchanged by this treatment variation. Therefore, the observability of the investment decision should not affect sellers' behavior. However, it might change how buyers select a seller. If they are attracted by sellers who have invested, sellers become more willing to invest.

[6] Buyers know that 20 is the highest possible valuation, but they do not know if they have a valuation of 20 for any version in the current round. When their inferred valuation is 20, they know that they have purchased the optimal variant. An inferred valuation of 19 is at least the second highest valuation and probably the highest. The expected value for their highest valuation is 19. If they derive valuations which lie on average below 19, buyers expect that they have not purchased their preferred variants in the past rounds. If the inferred valuation lies below 14, they know for sure that the recommendation has not maximized their payoff, because 14 is the highest valuation in the worst possible draw of valuations.

When watchdogs are present and sellers fear that each of their recommendations might be tested and graded with some probability, they might feel more inclined to invest in order to give better recommendations and receive better grades. Instead of giving selfish advice, sellers might refrain from giving advice when they have not invested. The increase in the quality of recommendations may potentially be reflected in higher prices. Introducing watchdogs may be a possibility to include more information in the market and facilitate reputation building. Of course, watchdogs can only affect the sellers' decisions if buyers react to the grades. Yet the effect of grades may wear off after some rounds: consumers might avoid badly rated firms for some rounds, but might give these firms another chance after some time. It might also be the case that consumers still pay more attention to the prices than to the sellers' grades such that the introduction of watchdogs does not change the sellers' pricing and investment decisions.

In the baseline treatment, I expect almost no investment decisions, very low prices and either selfish or no recommendations from sellers. In naturally occurring settings a market resembling the baseline treatment would consist of cheap discounters who do not offer reliable consumer service. The visibility of the investment decision might induce sellers to invest and give recommendations. However, it might also be the case that sellers refrain from investing because they do not want to know the impact of their recommendation on the buyers' payoff. They might find it easier to give a selfish recommendation when the buyers' consequences of this decision remain unknown to them. This reasoning would stand in line with the literature on moral wiggle rooms (see Dana et al., 2007).

While in the treatments without watchdogs only lying-averse informed sellers take into account the consumer's valuations, the threat of a visit by watchdogs might improve the consumer-friendliness of the recommendations.[7] Thus in the treatment with both observable investment decisions and watchdogs, the frequency and quality of the advice should be highest. However, this market setting might also provoke the coexistence of discounters who provide cheap goods, but no valuable advice, and quality stores that are expensive, but have good service. Therefore I am curious about the welfare consequences of the treatments: can watchdogs help overcome the problems of credence goods markets, i.e. do they induce sellers to give better recommendations? Is the visibility of the investment decision necessary to encourage sellers to invest? However, I am also interested if more investment and better recommendations actually increase

[7] There is experimental evidence that decision makers experience aversion to lying, e.g. Lundquist et al. (2009).

buyer welfare: the increase in welfare due to more purchases of the individually preferred versions may be offset by a raise in prices.

2.5 Results

Unless explicitly stated otherwise, I focus my analysis on rounds 11 to 25. Excluding the first rounds, I ensure that the treatment effects are not polluted by potentially erratic choices at the beginning of the experiment.[8] Similarly, I do not include the last five rounds (rounds 26-30) to rule out possible end game effects. Reported p-values are based on two-tailed two-sample or two-tailed one-sample Fisher Pitman permutation tests.

Table 2.2 provides a rough overview of the main decisions for each treatment. Posted prices are low in all treatments, yet they exceed 5 ECUs – the price at which a seller who has not invested can expect to break even. Consumers are very price sensitive, but less so in the O + W treatment. In the treatments without observable investment decisions sellers are reluctant to invest. Independent of the seller's investment decisions buyers usually receive a recommendation in all treatments before they select a version. About half of the recommendations are implemented in all treatments but the O + W treatment where almost three fourth of the buyers follow the advice. The share of buyer-friendly recommendations is maximal in the O + W treatment, but rather low on average.

Table 2.2: Overview of average results

	Baseline	**O**	**W**	**O + W**
Number of sessions	6	6	5	5
Posted price	6.24	6.15	6.62	6.43
Transaction price	4.43	4.88	4.73	5.48
Share of investing sellers	0.12	0.38	0.15	0.51
Share of purchase decisions with advice	0.97	0.91	0.84	0.93
Share of buyers following advice	0.50	0.50	0.45	0.74
Share of recommendations deserving a '1'	0.22	0.24	0.26	0.36

[8] Thereby I also eliminate observations in which a seller failed to set a price.

2.5.1 Recommendations

Do buyers follow the recommendations?

Starting with the last decision in a round and proceeding backwards, I first analyze how buyers interpret the recommendations they receive. Do they follow the sellers' advice? And how does their behavior depend on the sellers' investment decisions when they are observable? If they *know* that their seller has not invested, buyers should be indifferent between all five options as valuations and costs are uncorrelated and the recommendation contains no information about their most preferred version. If they *doubt* that their seller has invested or if they believe the seller wants to maximize his own payoff, they should also be indifferent between the recommended purchase decision and any other purchase decision. They should only favor following the advice if they believe that they face an informed and honest recommendation. If they assume that their seller has given a selfish recommendation and they want to stand in the way of the seller's profit maximization, they might choose any option but the recommended.

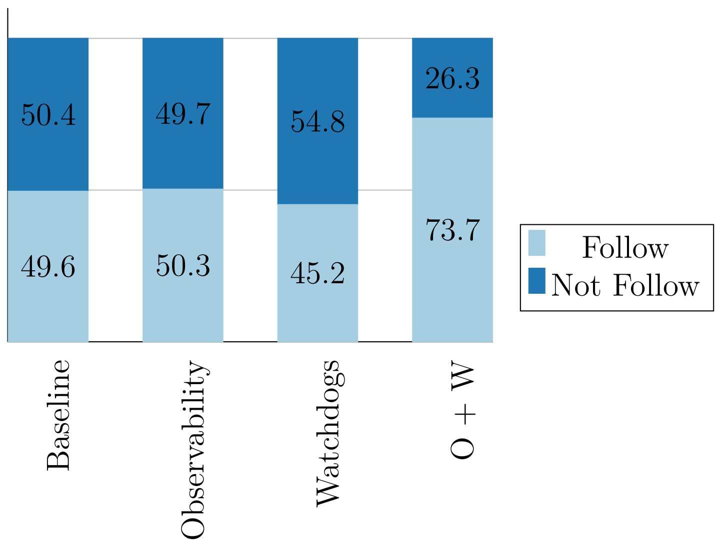


Figure 2.1: Percentage of buyers following the recommendation

As can be seen from Figure 2.1 buyers follow the advice in about 50% of the situations in all treatments except for the treatment with observable investment decisions and watchdogs, 'O+W'. In this treatment, almost 74% of the recommendations are implemented which constitutes a significant difference ($p = 0.023$) to all other treatments.

In order to be able to control for the price I regress the propensity to follow the recommendation on a dummy indicating whether watchdogs can occur, a gender dummy

Table 2.3: Regressions: Propensity to follow the seller's advice

	(1) Observability	(2) No Observability
Investment	0.226***	
	(0.062)	
Watchdogs	0.163**	-0.049
	(0.082)	(0.100)
Price	0.054***	0.029**
	(0.016)	(0.015)
Male	-0.045	0.022
	(0.080)	(0.100)
Period	0.030	0.037
	(0.036)	(0.035)
Period^2	-0.001	-0.001
	(0.001)	(0.001)
Constant	-0.169	0.086
	(0.308)	(0.343)
R^2	0.084	0.018
N	632	643

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and the price while controlling for the round (see Table 2.3).[9] I also include a dummy for the seller's investment decision as an explanatory variable in Regression 1. I do not include the seller's investment decision in Regression 2 because it remains unknown to the buyer such that she cannot base her purchase decision on it. I find the investment decision of the interacting seller, the presence of watchdogs and the price to have a positive and significant effect in the treatment with observable investment decisions. In the treatments in which the investment decision remains hidden, only the price has a positive and significant impact on the buyers' propensity to follow the advice.

Result 1: Buyers implement half of the recommendations they receive. When watchdogs are present and the investment decision is revealed, they follow almost 75% of the recommendations.

Figure 2.2 shows how the purchases depend on the observed investment decisions. When buyers know that their seller has invested, they follow his advice in 70% of the cases in the treatment with observable investment decisions although the sellers' incentives to give a consumer-friendly recommendation do not differ from the baseline

[9] I cluster the standard errors – which are presented below the coefficients – on seller level. In order to keep the interpretation of the coefficients simple, I present the outcome of a random effects panel regression. Other specifications yield very similar results.

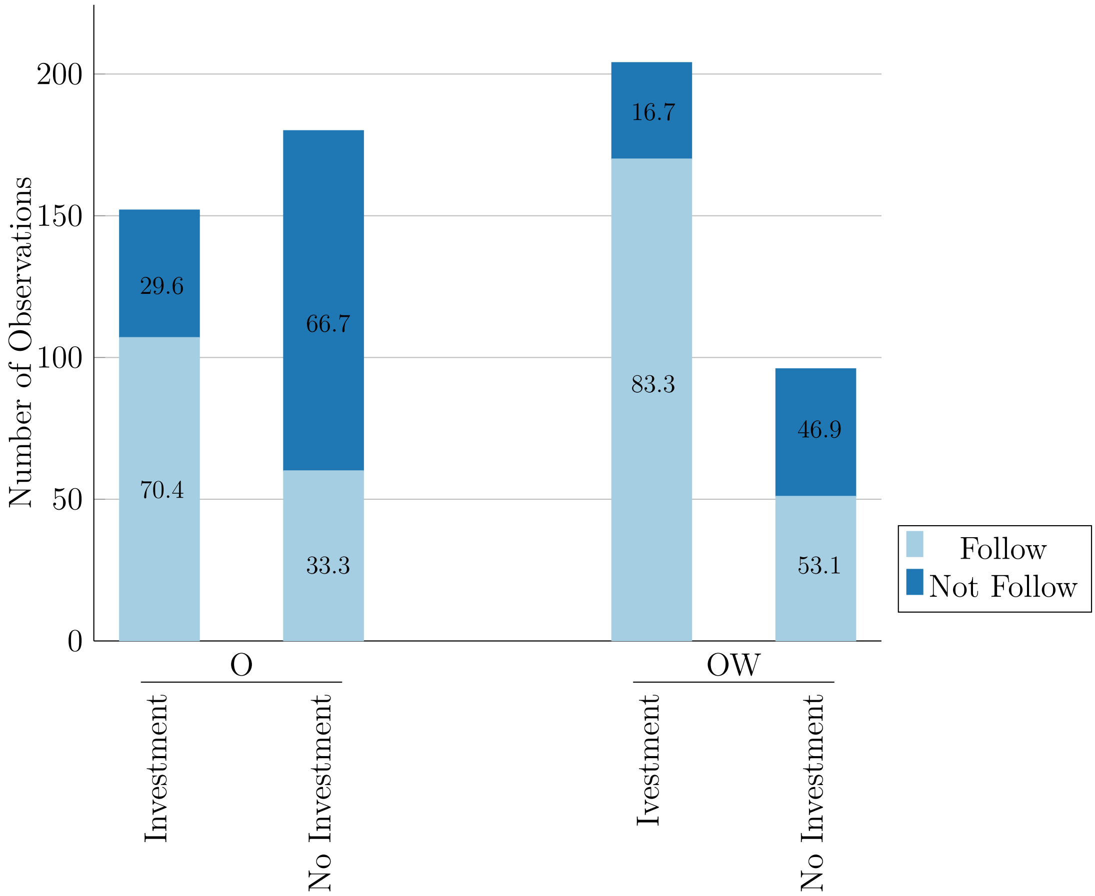


Figure 2.2: Implemented recommendations relative to the seller's investment

treatment. Thus, buyers seem to be confident that their seller will give a benevolent recommendation when he observes the consequences of his recommendations on their payoff. Whether sellers actually give consumer-friendly recommendations when they have invested will be analyzed in Section 2.5.1. When in addition to the observability also watchdogs can occur, the advice is implemented in 83% of the situations such that on average, 78% of the recommendations are implemented when sellers have verifiably invested.

Knowing that their seller has not spent 1.5 ECUs in order to observe their valuation and give a consumer-friendly recommendation 40% of the buyers still follow the advice – 33% in the treatment with observable investment decisions, 53% when additionally watchdogs are present. Hence, an investment almost doubles the rate with which buyers behave in line with the recommendation in the treatments with visible investment decisions.[10] However, the percentage of buyers following the obviously unqualified advice is surprising: in the treatment with observable investment decisions and watchdogs, more

[10]The difference is significant for the 'Observability' treatment ($p = 0.031$), but not for the 'O + W' treatment ($p = 0.125$).

Table 2.4: Average buyer valuations relative to following the recommendation

	Baseline	O	W	W + O
Follow advice	15.13	15.14	15.82	15.95
Follow advice of informed seller		17.41		16.49
Follow advice of uninformed seller		15.46		14.72
Do not follow advice	15.06	15.07	14.97	14.67

than half of the buyers follow the seller's advice although he has not invested. One reason for this behavior might be that buyers are efficiency-concerned: as prices are low, they know that sellers do not earn much. They might assume that the seller signals his profit maximizing variants through the recommendation. Since their expected valuation does not differ between the recommended version and the other four versions, they follow the seller's recommendation and thus implement an efficient outcome.

Do buyers benefit from following the advice?

As valuations and costs are uncorrelated even the most selfish recommendation of a seller does not necessarily harm buyers. Table 2.4 summarizes the average valuations buyers could have earned if they had followed the seller's advice. The table depicts the gross valuations, i.e. prices have not been subtracted. Furthermore, it also shows the mean of the expected valuations from choosing any option but the recommended one. From the table it appears as though it can never be harmful for buyers to follow the advice. Even when they do not know if the seller has invested (in the baseline treatment and in the treatment with watchdogs, 'W'), they are not worse off when they purchase the recommended option ($p = 0.875$ in the baseline treatment and $p = 0.125$ in the watchdogs treatment). In the treatments with observable investment decisions, buyers are clearly worse off when they follow the advice of an uninformed seller compared to following the advice of an informed seller ($p = 0.010$). Thus, in the treatments without observability it does not matter for their payoff if buyers follow the recommendations, yet in the treatments with observability, they should follow the advice of sellers who have invested.
Result 2: Buyers benefit from following the advice of informed sellers when the investment decision is observable.

When do sellers give recommendations?

In this section I want to investigate what causes sellers to give or not to give advice. Do they refrain from giving advice when they have not invested? And does the treatment influence their decision to give advice? In the vast majority of the purchase decisions, buyers receive a recommendation (91.45%). Figure 2.3 illustrates the four possible constellations of advice and investment for all purchase decisions in each treatment. The segments of the bars depict the share of purchase situations in which buyers received a recommendation and in which buyers did not receive a recommendation. Furthermore, it shows whether sellers who gave or did not give a recommendation invested. From Figure 2.3 it becomes clear that the propensity to receive a recommendation differs only slightly between the treatments: in the baseline treatment advice is provided in 96.87% of the purchase decisions, whereas in the treatment with watchdogs the percentage is lowest with sellers recommending a version in only 84.25% of the cases. Comparing the share of purchase situations in which advice is received with and without watchdogs I find that watchdogs have a mildly significant negative impact ($p = 0.067$) on the sellers' tendency to give a recommendation. This variation might stem from the fact that sellers who have not invested fear to get caught giving bad advice.

As expected, in the majority of situations when buyers do not receive a recommendation the corresponding seller has not invested. But the majority of the recommendations also stems from sellers who have not invested. Only in the treatment with watchdogs and observable investment decisions buyers receive informed advice in the majority of the purchase situations (64.7%). Thus, in this treatment 69.53% of all recommendations are given by sellers who can observe the buyer's preferences. Comparing the share of informed decisions when advice is given and when no advice is given, it becomes clear that the sellers' investment decisions do not have a significant influence on their choice to give a recommendation. Thus, also informed sellers might refrain from giving advice. A potential explanation for this observation might be a conflict between the seller's own and the buyer's interest, i.e. the version with the lowest costs for the seller leads to a low consumer surplus. If they recommend the option that maximizes their own payoff, they risk receiving a bad grade. Yet recommending the option leading to the highest buyer surplus could reduce their own profits and might even lead to losses. Consequently, they might prefer not to give advice and thus avoid taking responsibility for the outcome of the round. Whether this reasoning applies will be addressed in Section 2.5.1.

What do sellers recommend?

Focusing on the sellers who have invested, I want to find out which option they recommend. Figure 2.4 illustrates which share of all recommendations in a treatment is

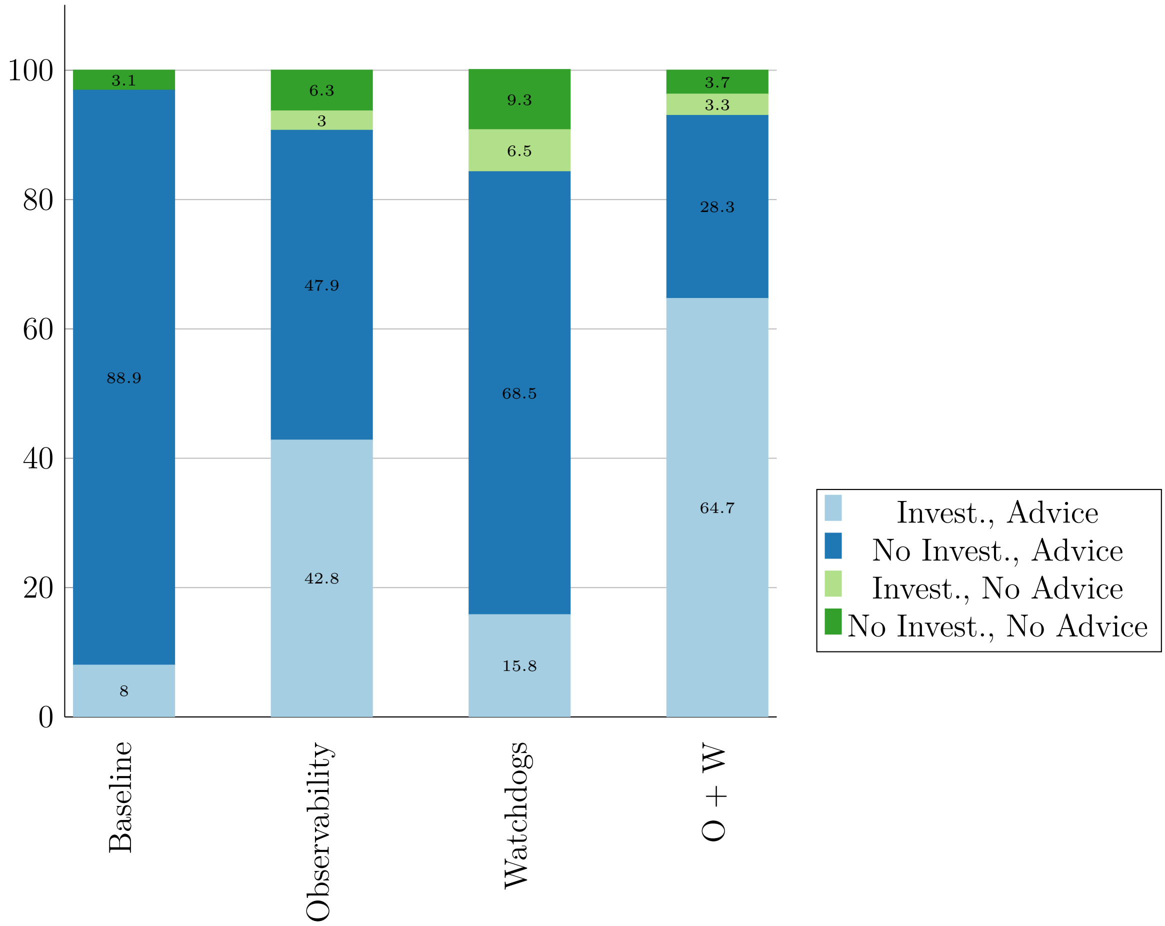


Figure 2.3: Advice relative to the seller's investment decision

given by uninformed sellers (right) and which share is given by informed sellers (left).[11] Furthermore, it depicts what percentage of the informed recommendations stand in line with the most obvious objectives: how often do sellers give selfish recommendations (S)? What percentage of the recommendations maximizes buyers' surplus (C)? And how often is it the sellers' aim to maximize joint welfare (J)? Furthermore, how often do they recommend options that fulfill several aims (CSJ, SJ, CJ) at the same time? The segments of the respective left bars show what share of the recommendations falls into each category. The right bar is not split up into different segments because sellers only observed their marginal costs and thus could not assess the impact of their valuation on their client's surplus.

In the baseline treatment only eight percent of the recommendations stem from sellers who have invested and of these informed recommendations almost all maximize the seller's profit (CSJ, SJ, S). Since valuations and costs are uncorrelated, these

[11]For each treatment, the two bars sum up to 100%, i.e. to all recommendations which were given in the respective treatment. The figure does *not* address *whether* sellers give a recommendation.

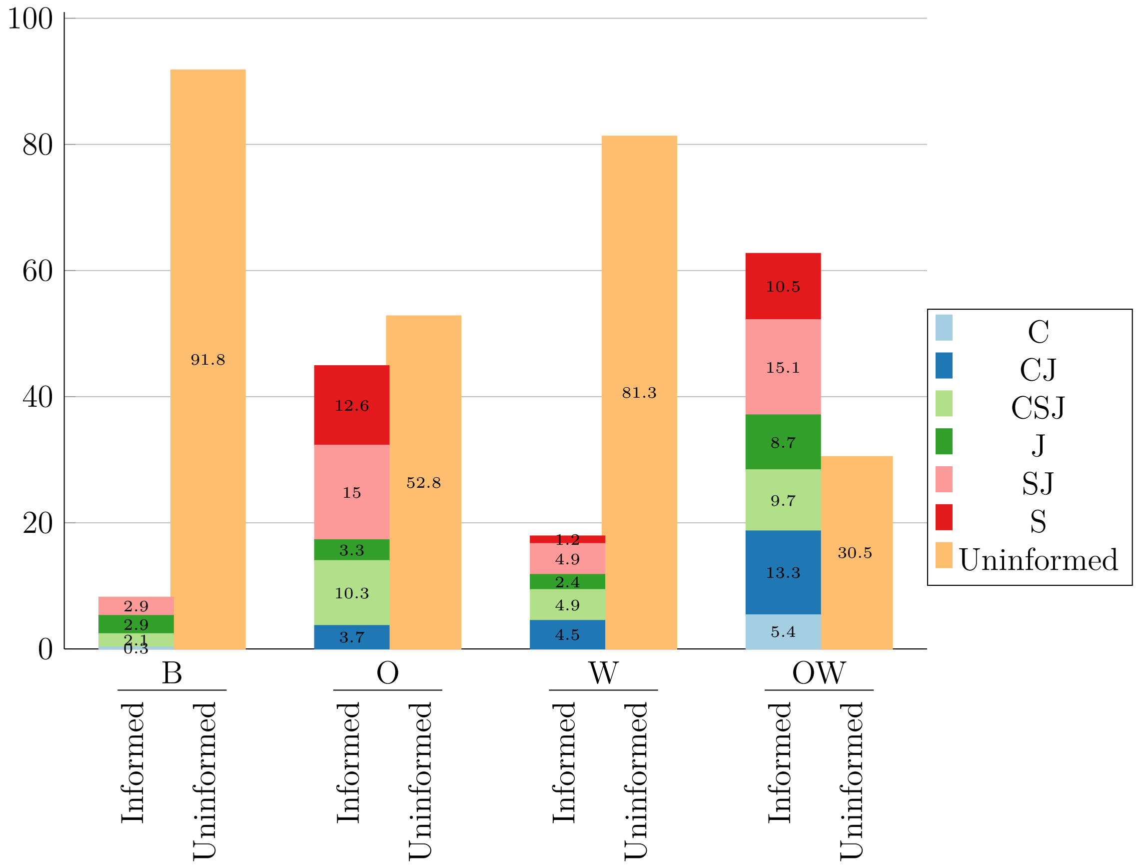


Figure 2.4: Content of the recommendations

recommendations happen to maximize consumer surplus in roughly one third of the situations (CSJ).[12] In the baseline treatment sellers recommend the consumer-friendly option only when it also maximizes their own profit. There is only one out of 28 recommendations which maximizes consumer surplus and is not optimal for sellers (C). About 60% of the recommendations are welfare-optimal (CSJ, SJ). This share is not significantly different from the expected share of welfare-optimal recommendations (0.65) if sellers only maximized their own payoff ($p = 0.500$).

When the investment decision is observable, almost half of the recommendations come from sellers who have invested, but in only 30% of the cases sellers use the acquired knowledge about their buyers' valuations to maximize their buyers' welfare (CSJ, CJ). There is no recommendation which maximizes only consumer surplus, but does not take into account overall welfare. Hence, most sellers use their investment decision rather as a misleading signal to attract clients. It might also be the case that they abandon their initial intention to give a consumer friendly recommendation when they discover

[12] I cannot reject the null hypothesis that this share is equal to 20% which would occur if sellers only tried to minimize their own costs.

a strong conflict of interest between themselves and their buyer. 80% of the informed recommendations aim at maximizing sellers' profits (CSJ, SJ, S) and almost 70% maximize the sum of buyer and seller surplus (CJ, CSJ, J, SJ). Although significantly more recommendations stem from informed sellers compared to the baseline treatment[13], the relative frequency of the three objectives I consider for the sellers' decision which variant to recommend are practically unaffected by the revelation of the investment decision.

When watchdogs are introduced, a larger share of recommendations (18.70%) is given by sellers who are informed about the buyer's preferences compared to the baseline treatment[14], but compared to the treatment with observability, fewer recommendations stem from informed sellers. The pattern according to which sellers decide what to recommend appears to have changed: in this treatment, 50% of the informed recommendations maximize the buyers' surplus (CJ, CSJ) and only 58.7% of the informed recommendations are selfish (S, SJ, CSJ). With almost 90% the share of welfare-maximizing recommendations is largest in this treatment. The decrease in selfishness and the increase in the share of welfare-maximizing messages are both significant at the 10% level, whereas the change in the share of consumer friendly recommendations is insignificant.

When both watchdogs are present and buyers know which sellers have invested ('O+W') almost 70% of the recommendations come from informed sellers. With only 51% selfish recommendations, this treatment is the one in which informed sellers are most willing to give up their own profit in order to maximize buyers' surplus. About two thirds of the recommendations are welfare-optimal which is significantly less than in the treatment with watchdogs (p = 0.032), but not significantly different from the share in the baseline treatment (p = 0.897). The share of consumer friendly recommendations lies below the treatment with watchdogs only, but above the treatments without watchdogs at 40%. Overall, the visibility of the investment decision induces sellers to invest more often and the presence of watchdogs motivates them to behave less egoistically and also take into account buyers' surplus.

Result 4: When the investment decision is observable, a larger share of the recommendations stems from informed sellers. The presence of watchdogs reduces the selfishness of the recommendations.

[13]The increase in the share of informed recommendations compared to the baseline treatment is significant at the 1%-level.

[14]This difference is only marginally significant with a p-value of 0.093.

Table 2.5: Percentage of sellers per grade in test periods

Grades	W	O + W	Baseline	O	W	O + W
Highest – 1	13.8%	32.5%	22.2%	23.6%	26.7%	34.9%
Mediocre – 3	65.5%	55.0%	59.2%	57.9%	55.7%	54.4%
Lowest – 5	20.7%	12.5%	18.7%	18.5%	17.7%	10.7%
	Awarded		**Hypothetical**			

Watchdogs are introduced to add more information to the market: in contrast to buyers, watchdogs can find out whether the recommendation was favorable to the buyer. I therefore analyze the awarded grades in the 'W' and in the 'O+W' treatment. Keep in mind that watchdogs do not differentiate between informed and uninformed recommendations. To avoid losing any observations on watchdogs I am using all 30 rounds for this part of the analysis. As can be seen from Table 2.5 most sellers received a '3' for giving a recommendation that was neither best nor worst for the buyer. The best grade was awarded more often in the treatment with observable investment decisions, for the other grades the differences are less pronounced. A Fisher's exact test reports no significant differences between the two treatments ($p = 0.178$). How buyers and sellers respond to the awarded grades will be analyzed in Section 2.5.2 and 2.5.3.

In addition, it might be interesting to find out what grades sellers could have achieved in the other treatments and in the rounds that were not tested. I therefore examine all purchase decisions and award a hypothetical grade '1' if the recommendation maximized the buyers' welfare, a hypothetical '5' if the advice minimized the consumer's welfare and a hypothetical '3' for all recommendations that were neither consumer-friendly nor particularly consumer-unfriendly. The percentages of recommendations falling into each category are also depicted in Table 2.5. If sellers determined their recommendations randomly or based on their (randomly determined) costs, the frequency of a '1' or '5' should not differ from 20% and a '3' should be awarded in 60% of the cases. According to one-sample χ^2-tests, the frequencies of the hypothetical grades do not differ significantly from the expected frequencies in the baseline treatment ($p = 0.334$). Hence, we cannot reject the null hypothesis that sellers do not take the buyers' valuations into account and determine the recommendations selfishly or randomly. When the investment decision is revealed, the difference in the distributions of observed and expected hypothetical grades is mildly significant ($p = 0.086$), yet in the two treatments with watchdogs the distribution of the hypothetical grades differs significantly from the expected distribution ($p = 0.009$ in the watchdogs treatment and $p < 0.001$ in the O+W treatment). Thus, the presence of watchdogs and the revelation of the investment decision induce sellers to give more consumer–friendly and fewer consumer-unfriendly recommendations.

What do sellers recommend when their interest and their client's interest are not aligned?

In situations with a trade-off between seller's profit and buyer's surplus, most sellers do not refrain from giving advice. Since only informed sellers can become aware of this conflict of interest, the following analysis deals only with sellers who have invested. I refer to situations as conflict situations when the buyer's most preferred option equals the seller's least preferred option. As illustrated in Figure 2.5, there is no informed seller in the baseline treatment who does not give a recommendation in such a conflict situation. In both treatments with observable investment decisions, there are few conflict situations in which sellers avoid giving recommendations. In the treatment with only watchdogs the probability of receiving a recommendation is lowest: in 24% of the cases sellers refrain from giving advice.

When watchdogs are absent, the majority of the recommendations are selfish (SJ, S). Under the threat of checks by watchdogs, sellers behave significantly less egoistically ($p = 0.008$) such that less than 50% of the recommendations maximize sellers' profit. The share of consumer-friendly recommendations (C, CJ) in conflict situations is negligible in the treatments without watchdogs and amounts to slightly less than one third in the treatments with watchdogs. In the baseline treatment the share of buyers receiving joint-welfare-maximizing messages is lowest (SJ) and it is highest in the treatments when watchdogs are present but the investment decision remains hidden (SJ, CJ, J). Overall, the presence of watchdogs provokes sellers to give welfare maximizing recommendations more frequently ($p = 0.029$) than in the treatments without watchdogs.
Result 6: The presence of watchdogs induces sellers to take into account consumer welfare by giving more consumer-friendly and fewer selfish recommendations in conflict situations.

2.5.2 Purchase Decisions

How do buyers select a seller?

Buyers can base their purchase decisions on their own experience with the sellers and on the sellers' grades, their investment decisions and their prices. First, I want to find out whether buyers reward sellers for high payoffs by purchasing from the same seller

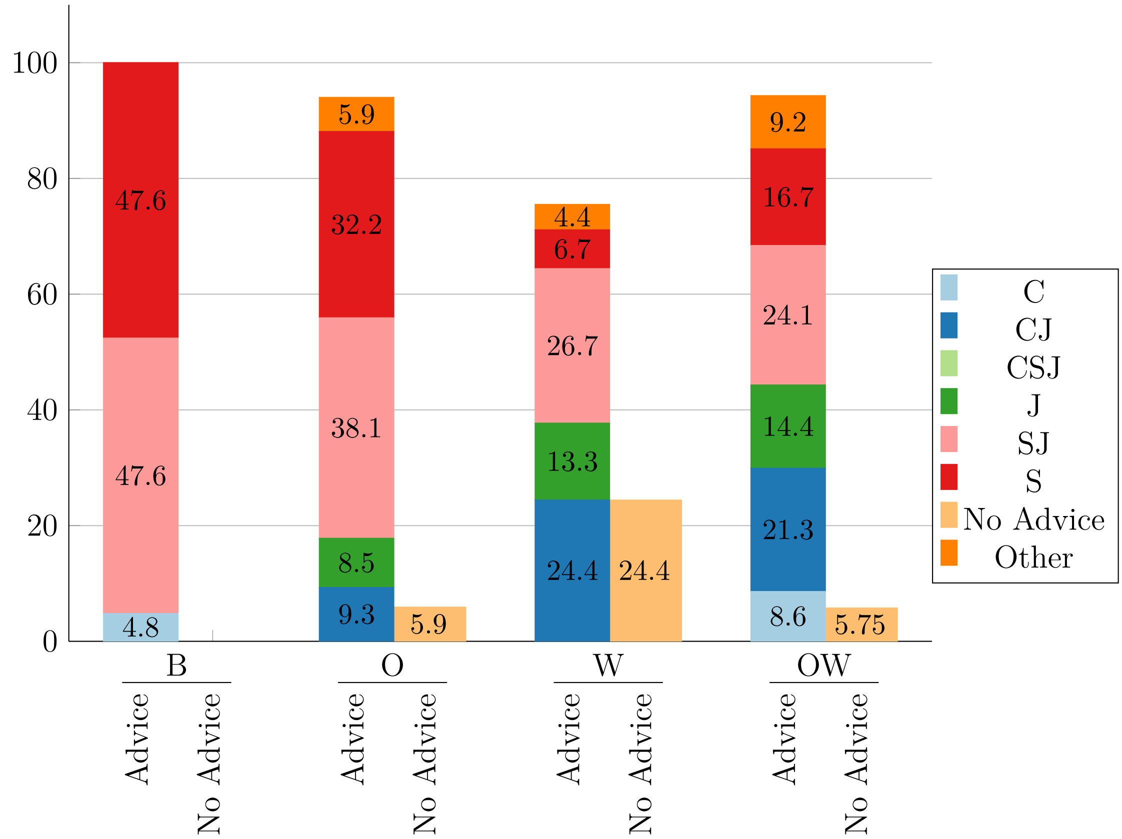


Figure 2.5: Recommendations in conflict situations by treatment

repeatedly. In fact, buyers are loyal if they earned a relatively high payoff in the previous round, but they tend to switch to another seller if their previous payoff was relatively low. Thus, the average payoff preceding a switch is significantly lower than the average payoff which is followed by the decision to purchase from the same seller again ($p < 0.001$). Hence, they punish sellers when they are unsatisfied with their payoff although they cannot always be sure whether the low payoff resulted from the seller's selfish or uninformed recommendation or whether they did not have higher valuations for the product versions in the last round.

How do buyers interpret the prices? Are they very price-sensitive or do they interpret higher prices as signal for the seller's investment? Figure 2.6 illustrates the share of buyers purchasing from the cheapest seller and Figure 2.7 depicts the propensity to purchase from the cheapest sellers for each treatment over time. In all treatments the majority of buyers purchases from the seller(s) setting the lowest price. In the baseline treatment, buyers decide in favor of the cheapest seller in almost 90% of all purchase decisions. When watchdogs are present or the investment decision is observable buyers seek the seller with the lowest price in 85% and 80% of the purchase decisions respectively. For these treatments the null hypothesis that the treatment does not influence the propensity to seek the cheapest seller compared to the baseline treatment cannot be rejected. In the treatment with watchdogs and a transparent investment decision however, buyers are significantly less price sensitive ($p < 0.001$) and do not purchase from the cheapest sellers in 42% of all cases.

Result 7: Buyers prefer cheaper sellers, but become less price-sensitive when both the investment decision is revealed and watchdogs are present.

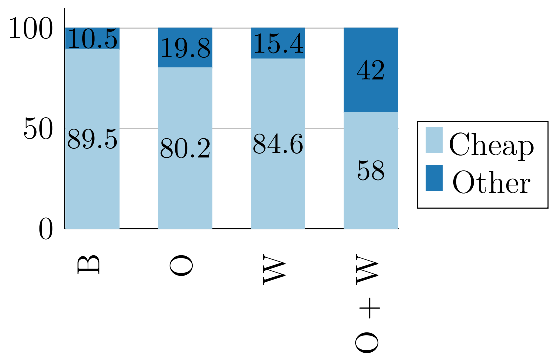


Figure 2.6: Share of buyers purchasing from the cheapest seller

It might be the case that buyers differ in their heuristics of choosing a seller, i.e. there might be some buyers who always select the cheapest seller whereas others never purchase from the seller with the lowest price. For this reason I show the distribution of purchases from the cheapest seller in Figure 2.8. In the baseline treatment, the treatment with observable investment decisions and the watchdogs treatment about 80% of the buyers choose the cheapest seller in almost all purchase situations. Only a small share of buyers deviates from this behavior by also occasionally purchasing from more expensive sellers in these treatments. The pattern in the treatment with watchdogs and observable investment decisions is different: here, only 25% of the buyers always choose the cheapest seller while other buyers have a less pronounced preference for the cheapest seller and frequently purchase from more expensive sellers.

Do buyers prefer sellers who have invested?

The sellers' intentions to invest can be twofold: on the one hand, sellers might invest to be able to give better advice while on the other hand, they might only use their observable investment as a signal to attract buyers. I therefore seek to find out whether buyers take the sellers' investment decision into account. When at least one seller has invested and buyers can observe the outcome of the investment decision, the majority of the buyers selects a seller who has invested. In the treatment without watchdogs 52% of the buyers interact with a seller who has invested when given the choice, while in the O

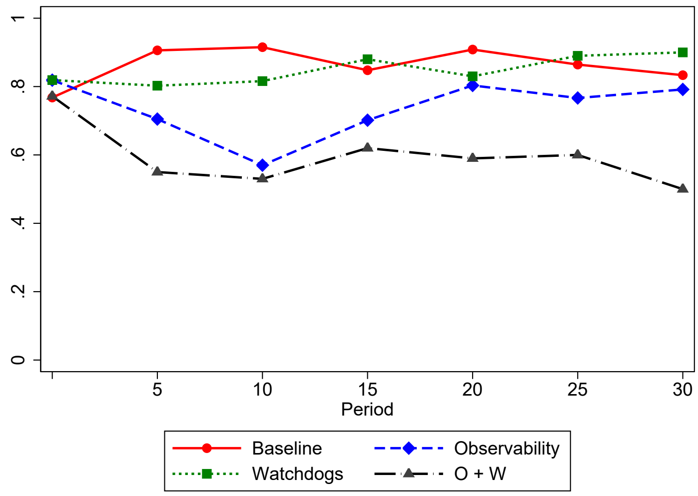


Figure 2.7: Propensity to choose the cheapest seller

+ W treatment 71% of the buyers choose a seller with a positive investment decision.[15]

As there may again be differences between buyers when interpreting the sellers' investment effort, I also take a closer look at the distribution of individual purchases from sellers who have invested. As can be seen from Figure 2.9, preferences for sellers who have invested differ a lot among buyers in the Observability treatment: some buyers purchase almost never from a seller who has invested, others only purchase from sellers who have invested and there are also undecided buyers who sometimes choose sellers who have invested. In contrast, in the treatment with watchdogs and observable investment decisions the majority of buyers always seeks sellers who have invested. In both treatments, there is no systematic difference in the average prices paid by those who usually purchase from sellers who have invested and those who normally purchase from sellers who have not invested.

Do buyers prefer sellers who have received a good grade?

Buyers cannot tell whether the *current* round is a test round, but they are informed if watchdogs have been present in the past and which grades have been awarded to the four sellers in their market. One could imagine that buyers react to the grades by seeking sellers who have received a good grade as they might expect these sellers to also give good recommendations in the current round. Figure 2.10 depicts the average number of buyers that sellers were able to attract in the test round and in the four rounds after

[15]For a seller-focused analysis of the investment decision see Section 2.5.4.

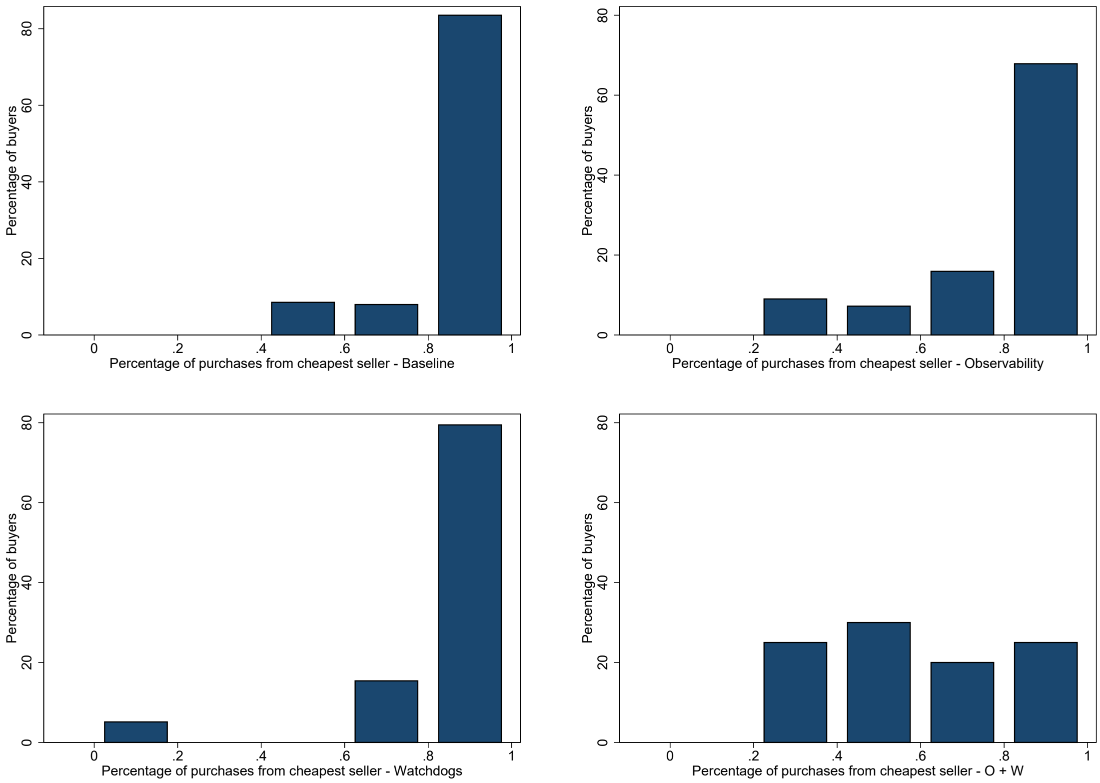


Figure 2.8: Distribution of individual purchase decisions: cheap

the appearance of watchdogs.[16] In the figure, I only consider markets in which different grades were awarded. There were no markets in which a ‘1’, a ‘3’ and a ‘5’ were awarded in the same round. Additionally, I combine the three combinations of grades in the figure by only referring to the relatively better and the worse grade. When a seller is awarded a ‘3’ and another seller in the market receives a ‘5’, the clients of the seller with the ‘3’ are described as the clients purchasing from the seller with the better grade. From Figure 2.10 it becomes clear that the above mentioned reasoning is not very widespread among buyers: even in the first round following the test round sellers with better grades are not chosen by significantly more buyers than their competitors with worse grades ($p = 0.449$) and also afterwards buyers are not discouraged to select a seller with a worse grade. An explanation for this surprising fact might be the low price level which ensures a relatively high consumer surplus even when random purchase decisions are made.

[16]I combine all observations on test rounds and set the test round equal to zero. Again, I am using all rounds of the experiment.

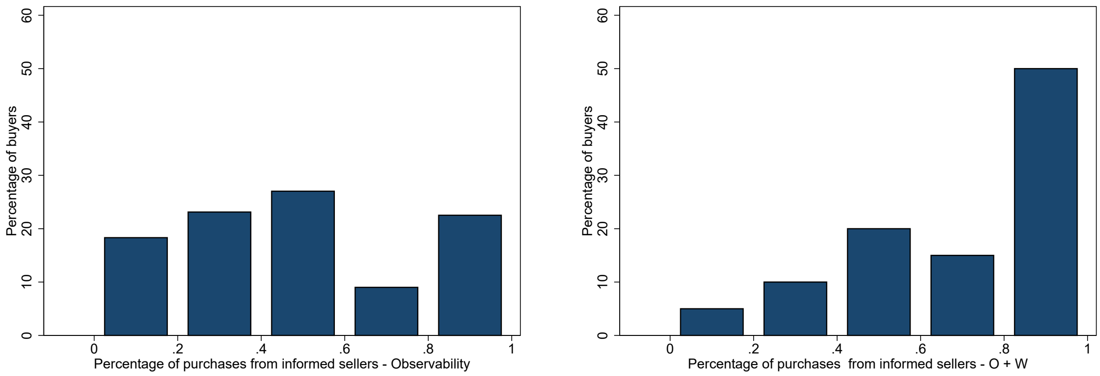


Figure 2.9: Distribution of individual purchase decisions: investment

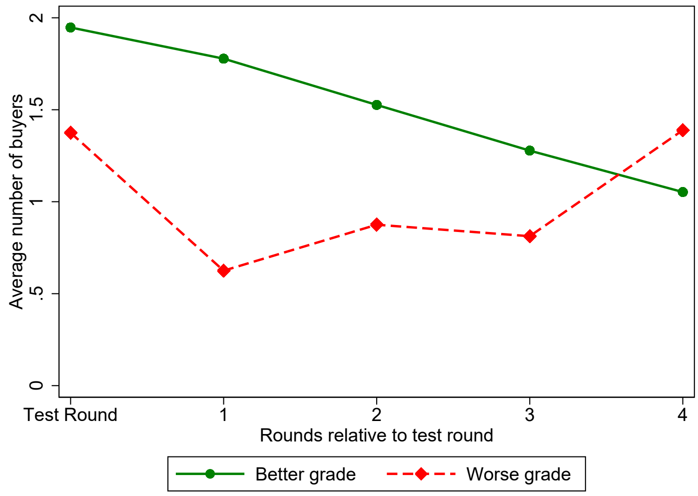
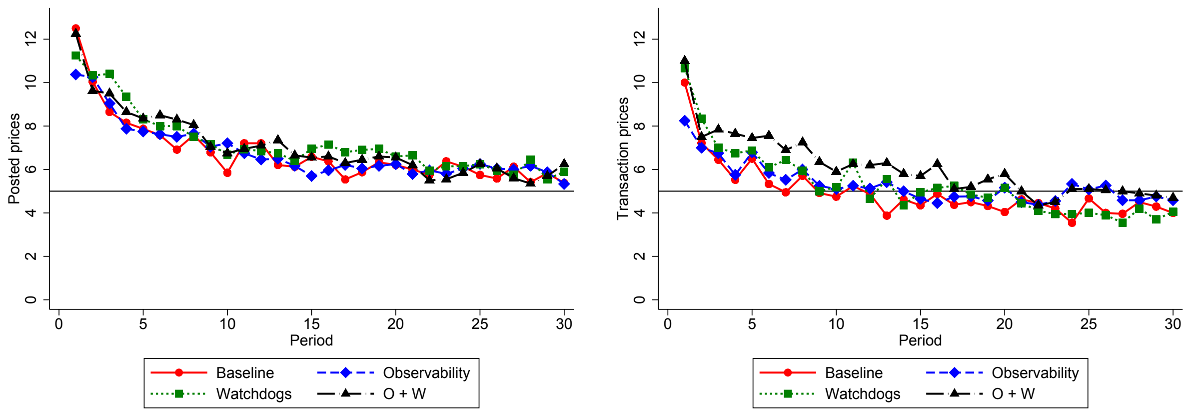


Figure 2.10: Average number of buyers per seller relative to the awarded grade

2.5.3 Prices

What prices do sellers set?

Having analyzed how buyers select a seller, I now turn towards the sellers' price setting decision. In all treatments sellers charge relatively high prices at the beginning of the session. Prices sometimes even exceed a price of ten which would split the expected payoff equally. Realizing that lower prices attract more buyers, sellers adjust their price level (see Figure 2.11a). After about ten rounds, posted prices stabilize at six ECUs. In the rounds thereafter, the price level decreases very slowly and approaches slightly less than six ECUs towards the end of a session. As can be seen from the graph, the posted prices do not differ across the different treatments. The null hypothesis of identical posted prices in treatments with observable investment decisions and in treatments with unobservable investment decisions cannot be rejected (p-value is 0.747). Posted prices are not affected by the presence of watchdogs ($p = 0.423$).



(a) Posted prices **(b)** Transaction prices

Figure 2.11: Time series of prices

Result 8: There are no treatment effects on the posted prices. Posted prices are low, but not as low as 5 ECUs. Thus, prices exceed sellers' average costs.

Considering the transaction prices (see Figure 2.11b), i.e. the prices at which trade has actually taken place, a very similar pattern can be observed. However, the average transaction price is lower because buyers are very price sensitive. Transaction prices in treatments with unobservable investment decisions converge to 4 ECUs, whereas trade takes place at slightly higher prices in the treatments with observable investment decisions because buyers accept slightly higher prices when the investment decision is revealed. The increasing effect of this revelation on prices is marginally significant in a one-tailed two-sample Fisher-Pitman permutation test ($p = 0.069$) and thus insignificant in the two-tailed version of the test. From the graph it becomes clear that this effect is mainly driven by the higher transaction prices in the treatment with watchdogs and observable investment decisions. However, transaction prices when watchdogs are present and investment decisions are observable do not differ significantly from prices when the investment decisions are revealed, but watchdogs are absent ($p = 0.262$).

Result 9: Transaction prices are highest in the treatment with observable investment decisions and watchdogs.

Do sellers charge higher prices when they have invested?

Sellers might try to pass on the investment costs to buyers by adding a markup to their price. Charging a higher price they might also try to signal their investment decision to buyers when the investment decision is not revealed. Table 2.6 provides an overview of the average posted prices charged by sellers who have invested and those who have not invested in each treatment. In the treatment with an observable investment decision

Table 2.6: Prices and investment decisions

	Baseline	O	W	W + O
Posted prices				
Investment	6.57	5.93	5.80	6.45
No investment	6.20	6.28	6.77	6.41
Transaction prices				
Investment	4.57	4.89	4.63	5.72
No investment	4.42	4.87	4.75	4.97

(O) and in the treatment with watchdogs (W), prices are on average *lower* when the seller has invested.[17] In the baseline treatment and in the treatment with an observable investment decision and watchdogs (O+W), the average posted price with investment exceeds the average price without investment, but the difference is negligible. Thus, sellers do not use their price level to signal that they have invested in the treatments with hidden investment decisions.

Result 10: Sellers do not charge higher prices when they have invested.

To understand how buyers interpret the prices, I also take a closer look at the transaction prices. When buyers observe a positive investment decision they might be more willing to accept a higher price, because the seller is capable of giving a more valuable recommendation. The average transaction prices for each treatment are also summarized in Table 2.6. As expected there is no difference in transaction prices in the treatments with unobservable investment decisions. Yet in treatments with observable investment decisions, buyers accept higher prices from sellers who have invested ($p = 0.055$).

Do sellers react to their grade by altering their price?

At the end of a test round sellers are informed about their grade. Knowing that buyers observe their grades in the following rounds when they decide which seller to select, sellers might find it profitable to adjust their prices to their grades: sellers with relatively good grades might expect buyers to become less price-sensitive. Similarly, sellers who have received a bad grade might want to lower their price in order to attract more buyers. In Figure 2.12 I display the average price charged by buyers with different grades in the rounds after the test round. Sellers do not seem to react to their grade by adjusting their

[17] The result is marginally significant with a p-value of 0.063 in the treatment with watchdogs and insignificant ($p = 0.750$) in the treatment with observable investment decisions.

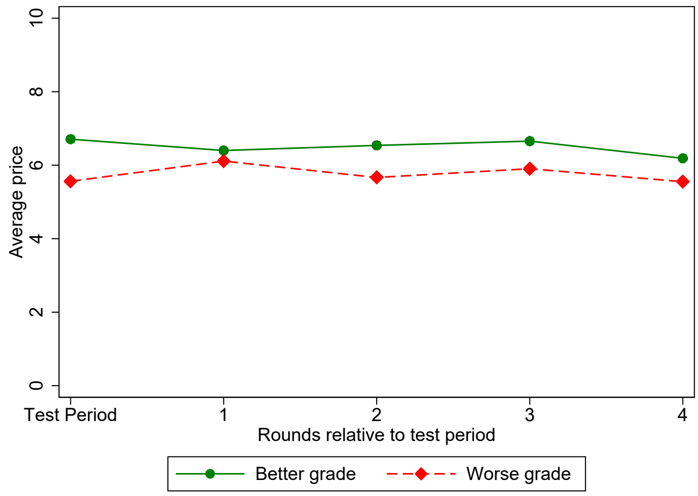


Figure 2.12: Average posted price for each grade

Table 2.7: Investment decisions

Setting	No Watchdogs	Watchdogs	p-value
No observability	12.2%	15.4%	*0.675*
Observability	37.9%	50.7%	*0.221*
p-value	*0.019*	*0.016*	

prices. The prices charged by sellers with better grades do not differ from prices charged by sellers who were awarded worse grades (p = 0.812). Since consumers do not react to the sellers' grades by avoiding those with bad grades as analyzed in Section 2.5.2, sellers are not induced to react by adjusting their prices.

2.5.4 Investment Decisions

Do sellers invest?

Investing seems unattractive for sellers because they cannot charge buyers for their investment costs directly nor can they set higher prices. Yet there is a substantial fraction of sellers who decide to invest as Figure 2.13 shows. The treatment does not play a role at the beginning of a session when the majority of sellers invests. Growing more experienced, some sellers cease to invest, but the investment rates do not decrease to 0 in any treatment.

In the baseline treatment, on average 12.2% of sellers invest and in the treatment with watchdogs but without observable investment decisions 15.4% of sellers invest (see Table 2.7). The effect of the introduction of watchdogs on the investment behavior in

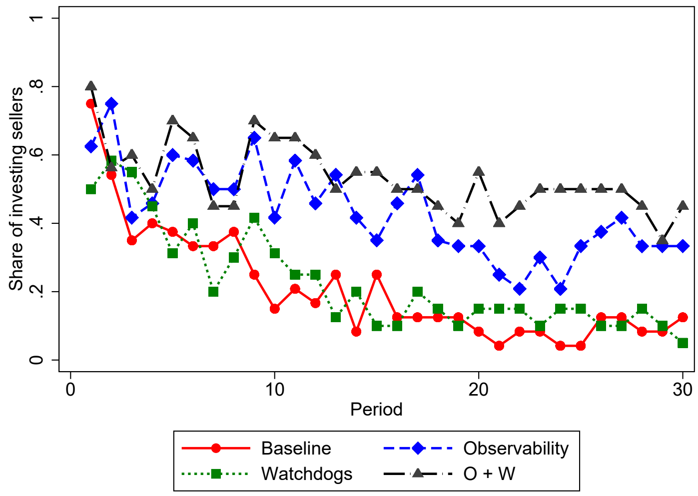


Figure 2.13: Time series of investment decisions

treatments with unobservable investment decisions is insignificant (p = 0.675). When buyers observe the investment decisions, but watchdogs are not present in the market, 37.9% of sellers pay in order to be able to observe the buyers' valuations, whereas in the treatment with observable investment decisions and watchdogs, 50.7% of sellers invest. This difference in the investment decision induced by the presence of watchdogs is again not statistically significant (p = 0.221). Pooling over the treatments with watchdogs and without watchdogs, the impact of the observability of the investment decisions is highly significant ($p < 0.001$): when buyers know which sellers have invested, sellers are more likely to invest. This observation stays significant when analyzing the treatments with watchdogs and without watchdogs separately.

However, there are large differences in the investment behavior between sellers (see Figure 2.14). In the baseline treatment and the treatment with watchdogs the majority of sellers never or almost never invests. In the treatments with observable investment decisions the pattern is different: there are almost as many sellers who always invest as there are sellers who never invest. The share of sellers who sometimes invest is considerably larger compared to the treatments without observable investment decisions.

Do sellers who invest attract more buyers?

Although sellers cannot demand higher prices when they have invested, they might still find it profitable to invest if it helps them attract more clients. If all sellers set the same price and investment decisions did not matter, each seller should be able to attract on average one buyer. With only 0.64 buyers per seller in the baseline treatment, sellers attract *fewer* buyers when they invest, but not significantly so (p=0.125). This observation

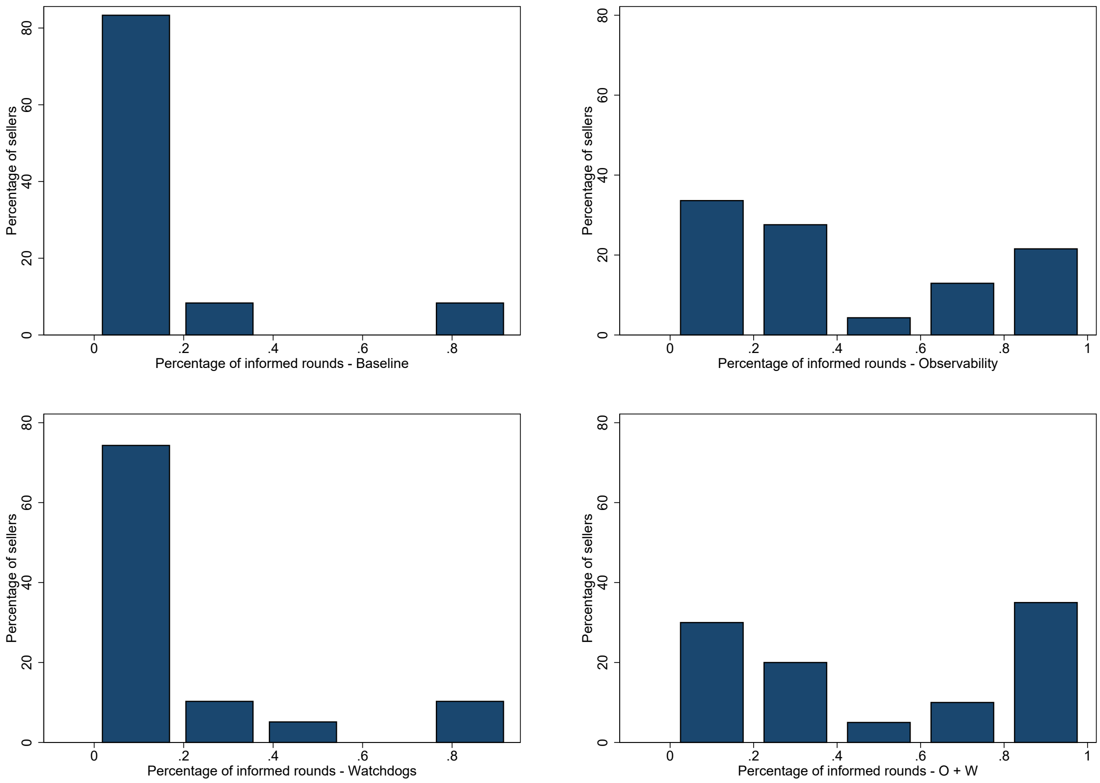


Figure 2.14: Distribution of individual investment decisions

Table 2.8: Clients per seller

	Baseline	**O**	**W**	**W + O**
Investment	0.64	1.15	1.49	1.20
No investment	1.02	0.83	0.92	0.89
p-value	0.125	0.438	0.313	0.125

might stem from the fact that in this treatment sellers who have invested post slightly higher prices and are thus less likely to find a buyer. Yet in the other three treatments (see Table 2.8), sellers who have invested attract more clients, but none of the effects is statistically significant at any conventional level. To analyze the effect of the investment decision on the number of attracted buyers I also perform a regression analysis which allows me to control for the price charged by the seller. According to the regression results (see Table 2.9), sellers who have invested attract significantly more buyers than their uninformed competitors in the treatments with observable investment decisions.[18]

[18]The difference is only significant in the regression analysis because the included control variables help to distinguish the effect of the investment decision from other potential influencing factors.

In both treatments, prices have a decreasing, but non-linear effect on the number of attracted clients. I find a negative coefficient for the interaction of prices and the investment decision. Thus, consumers are also price-sensitive towards informed sellers.

Table 2.9: Regression: Number of clients

	(1) Observability	(2) No Observability
Price	-0.833***	-1.050***
	(0.093)	(0.089)
Price^2	0.039***	0.047***
	(0.005)	(0.005)
Price * Investment	-0.091**	0.020
	(0.041)	(0.056)
Investment	0.920***	-0.214
	(0.340)	(0.509)
Watchdogs	0.080	0.203
	(0.188)	(0.131)
Period	-0.015	-0.021**
	(0.010)	(0.010)
Gender	-0.027	0.001
	(0.185)	(0.127)
Constant	4.611***	5.794***
	(0.454)	(0.412)
R^2	0.2419	0.4045
N	633	652

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Output from random effects panel regressions
Standard errors are clustered on seller-level

Result 12: Sellers who have invested and do not charge higher prices than their competitors attract more buyers in the treatments with observable investment decisions.

2.5.5 Welfare

In this section I seek to analyze whether the introduction of watchdogs and the revelation of the investment decisions lead to significant welfare consequences for buyers and sellers. In all four treatments sellers' payoffs are on average very low (see Table 2.10). This is due to the intense price competition which drives down prices. Besides, higher prices are not used or not understood as a signal for better recommendations. As prices decrease below the average costs of 5 ECUs and buyers prefer sellers who set very low prices, sellers incur losses when consumers select a variant with costs exceeding the price. Since low profits for sellers correspond to a higher payoff on the other market side, buyers benefit from the price competition in all four treatments equally by paying low prices. The treatment with observable investment decisions and watchdogs, intended to make buyers better off, actually maximizes *sellers'* surplus, because buyers accept higher prices in this treatment. However, neither for sellers nor for buyers are any of the differences in payoffs between the treatments statistically significant. For buyers the benefit from better recommendations is mostly offset by the premiums they pay for advice.

Table 2.10: Mean profits

	Baseline	**O**	**W**	**W + O**
Sellers	0.75	0.85	0.60	1.21
Buyers	10.72	10.12	10.72	10.32
Welfare	11.47	10.97	11.32	11.53

Table 2.11: Profits by investment

Sellers	**Baseline**	**O**	**W**	**W + O**
Investment	-0.91	1.09	-1.08	1.64
No investment	0.98	0.70	0.91	0.76
Buyers				
Investment	11.93	10.52	11.12	10.60
No investment	10.65	9.84	10.60	9.73

As summarized in Table 2.11 sellers are better off when they invest in the treatments with observable investment decisions, yet in the other two treatments they suffer significantly from investing, whereas buyers always benefit from interacting with a seller who has invested. Controlling for the price and experience the regression results (see Table 2.12) support this finding for the sellers, while the effects are less pronounced for buyers. Interestingly, prices have a non-linear effect on sellers' payoff: sellers benefit from small increases as they earn more from each transaction, but suffer from larger increases because fewer transactions take place.

Result 13: For buyers, the benefit from better recommendations due to the introduction of watchdogs and the transparency of the investment decision is

offset by an increase in prices.

Table 2.12: Regressions: Payoffs

	(1) Sellers	(2) Buyers
Price	0.976***	-0.289
	(0.286)	(0.270)
Price^2	-0.058***	-0.064**
	(0.016)	(0.025)
Investment*	-1.672***	0.554
	(0.590)	(0.413)
Observability	-0.006	-0.456*
	(0.341)	(0.242)
Investment * Observability	1.820**	0.496
	(0.773)	(0.496)
Watchdogs	0.002	0.349*
	(0.363)	(0.185)
Period	-0.076**	-0.039**
	(0.031)	(0.019)
Gender	-0.047	-0.117
	(0.362)	(0.182)
Constant	-1.203	14.046***
	(1.459)	(0.807)
R^2	0.0543	0.1872
N	1285	1275

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Output from random effects panel regressions
Standard errors are clustered on subject-level
*For buyers, investment refers to their interaction partner's investment.

2.6 Conclusion

This paper investigates the role of recommendations and costly investment decisions in horizontally differentiated credence goods markets. In these markets buyers suffer from very poor information: they can neither identify the product version they need before purchasing, nor do they find out if they have purchased their personally optimal variant ex post. Their informational disadvantage may be exploited by sellers. With the help of an experiment I try to understand the functioning of credence goods markets with a horizontal product differentiation component and identify potentially relevant influencing factors.

In the experiment subjects trade differentiated goods on markets taking the roles of buyers and sellers. Each seller in a market offers the same product variants as his competitors, but he incurs different marginal costs for each variant. He sets one price for all his variants. Consumers differ in their valuations for the product versions which are unknown to them. Consequently, they cannot identify their most preferred variant and would need to make a random purchase decision. Sellers can recommend a variant to their customer. However, they need to invest in order to be able to observe their clients' valuations. Otherwise, they only know their own costs for each variant which do not include any information about the consumers' preferences. Observing a recommendation, consumers do not know if their seller has invested and is thus *capable* of giving a useful recommendation. Even if they assume that their seller has invested, they do not know either if their seller is *willing* to take his potential knowledge about their preferences into account when deciding which version to recommend. Having followed his recommendation, they cannot judge its quality: observing their payoff they can derive their valuation for the purchased variant. Yet, they do not know for sure if they would have had a higher valuation for another variant. Since consumers will almost never find out if their sellers have given them reliable recommendations, sellers cannot build a reputation of giving good recommendations. To improve the market outcome in credence goods markets with horizontal product differentiation I implement two treatment variations: I reveal the sellers' investment decisions and I introduce watchdogs which test the quality of some randomly selected recommendations. Watchdogs are designed to resemble checks by consumer organizations while the revelation of the sellers' investment decisions comes close to the public display of qualification certificates.

In the experiment, consumers observe all investment decisions in their market before deciding with which seller to interact. In naturally occurring markets, consumer may need to visit a store and thus incur some search costs before they can gather information about the seller's qualification. However, searching for relevant information online is easy and sellers have a strong interest to show their qualifications. With watchdogs appearing in four rounds to judge and reveal the quality of sellers' recommendations, the frequency of their visits is most likely higher than the frequency of actual tests by consumer organizations. Yet, checks are also performed by magazines and online platforms such that the real frequency of tests might even exceed the one in the experiment in some markets. As most test results are published online, they are readily available to consumers.

In the baseline treatment with hidden investment decisions and no checks of the recommendation quality, very few sellers invest and these uninformed sellers give recommendations which aim at minimizing their costs. Even those who have invested give selfish recommendations. The price level is low as consumers are very price-sensitive. Although they might anticipate that their seller has not invested, they still follow about half of his recommendations. When the investment decision is revealed, more sellers decide to invest. Indeed, they rather invest to attract consumers than to base their recommendations on the acquired information as the share of selfish recommendations given by informed sellers is unchanged compared to the baseline treatment. Consumers do not show a clear preference for informed sellers, but if they choose to interact with an informed seller, they normally follow his recommendation. A non-negligible share of buyers even follows the recommendations of sellers who have obviously not invested. Knowing that each recommendation may be subject to a check by watchdogs, informed sellers give more consumer-friendly recommendations. The share of sellers who decide to invest is yet unchanged compared to the baseline treatment. Buyers are very price-sensitive in this treatment as well and they react little to the sellers' grades. In the fourth treatment, both the investment decisions are revealed and watchdogs appear from time to time. In this treatment, a substantial fraction of sellers invests and gives consumer-friendly recommendations. Buyers accept higher prices such that sellers who have invested do not incur losses in expectation. Also the fraction of buyers following the sellers' recommendations is highest in this treatment.

Thus, the situation in the baseline treatment is comparable to a market in which only discounters exist. Although they do not receive reliable advice, the low price level ensures that consumers still benefit from their purchases even if they happen to purchase their least preferred variant. The revelation of the investment decision induces more sellers to invest, but the quality of the advice is unchanged and thus the match between consumers' preferences and purchased variants is not improved. Watchdogs improve the quality of the recommendations, i.e. the share of expert sellers relative to discounters increases, but sellers are still reluctant to invest. Only when the investment decisions are revealed *and* watchdogs are present, sellers decide to invest and take into account their knowledge about buyers' preferences to determine their recommendation. In terms of welfare, the market constellations do not differ significantly. With both observable investment decisions and watchdogs matches are better, but also prices are higher compared to the other treatments. Hence, the simultaneous introduction of watchdogs and the revelation of the sellers' investment decisions provokes a switch from a situation with low prices and low-quality recommendation to a situation with higher prices and informative product recommendations.

For future research it might be interesting to change the parameterization of the experiment in order to investigate if the findings depend on the choice of parameters. The experiment might also be extended by increasing the frequency of visits by watchdogs or by introducing expert consumers who are informed about their valuations. In addition, allowing for consumer feedback might also have an impact on the market outcome. Treatments in which sellers do not compete in prices might yield new insights in the functioning of credence goods markets with horizontally differentiated products. Furthermore, I would like to introduce the possibility to gather second opinions.

Appendix

Instructions

Thank you for participating in this experiment. Please do not to talk to the other participants during the experiment.

This experiment consists of 30 rounds. There are two different roles you can take: sellers and buyers. Your role is determined randomly at the beginning of the experiment and you keep your role throughout the whole experiment. Which role you take is displayed on your screen.

The participants are split into two groups consisting of four sellers and four buyers each. These groups persist for the whole experiment. As a buyer your potential interaction partners are Seller 1, Seller 2, Seller 3 and Seller 4. If you are a seller, your potential interaction partners are Buyer 1, Buyer 2, Buyer 3 and Buyer 4. Note that the numbering of the sellers is fixed, i.e. the same number always describes the same participant. The numbering of the buyers is not fixed and changes from round to round.

In each round sellers offer the same five product variants for which different costs occur at the different sellers. These costs always lie between 0 and 10 ECUs. The sellers own enough units of each variant to serve all buyers in their group. Products and costs vary from round to round. Buyers have different valuations for the different variants. These valuations, which are randomly determined in each round and lie between 10 and 20 ECUs, are unknown to buyers. For this reason sellers have the possibility to give advice and assist buyers in choosing a variant. In order to know the buyers' valuations, a seller needs to invest in the beginning of a round. Investing costs 1.5 ECUs. A seller can only observe the buyers' valuations and give useful recommendations if he has invested. Buyers do not know if their seller has invested and if the recommendation is honest, i.e. if the seller has recommended the variant for which they have the highest valuation.

Payoffs in a round

Buyers

If a buyer has bought a product variant, her payoff equals her valuation for the chosen variant minus the paid price. If a buyer has not bought, her payoff in this round equals zero.

Sellers

If a seller has found *one* buyer, his payoff equals his price minus the costs of the sold variant minus his investment costs if he has invested.

If he has attracted *several* buyers, his payoff will be the sum of the payoffs from the individual transactions. The investment costs will only be subtracted once if he has invested.

If a seller has not attracted a buyer and he has not invested, his payoff equals zero. If has invested but no buyer has purchased from him, his payoff equals -1.5 ECUs.

Timing of a round

1. Investment decision: Sellers decide if they want to invest. For the investment costs of 1.5 ECUs occur.

2. Price setting: Each seller sets one price for his five product variants. This price needs to lie between 1 and 20 ECUs and applies to all five product variants. Each seller offers the same five product variants. When deciding on his price, the seller observes his costs for the five product variants. These costs occur when a variant is sold and they are independent of the investment costs. Note: If a seller does not set a price, the computer will save a price of 20 ECUs after 35 seconds. Thus, he will most likely not attract a buyer in this round.

3. Selection of a seller: Buyers observe the prices set by Seller 1, 2, 3 and 4 in their group. Buyers cannot identify the seller(s) who invested. Each buyer decides with which seller she wants to interact. If she does not select a seller, the round ends for her. As a reminder her past transactions are displayed on her screen.

4. Advice: Sellers get to know which buyers have chosen to interact with them. At most a seller can attract all four buyers from his group. If the seller has invested, he can observe the buyers' valuations for the different product variants. The valuations differ among the buyers. For each interacting buyer the net utility (= valuation - price) of each variant is displayed. The seller also observes his margin (= price - costs) for each product variant. If he has not invested, he only observes his margins. The seller decides for each buyer separately if he wants to give a recommendation and which variant to recommend.

5. Purchase: Buyers observe which variant their seller has recommended or that their seller has not given a recommendation. Buyers can follow the advice and purchase the recommended variant, but they can also choose another variant or decide not to purchase in the current round.

6. Payoff: The earnings are displayed on the screens. Sellers are informed if they could sell something and which variants their buyers' have chosen. If they have purchased, buyers are informed about the net utility (= valuation - price) of their purchase.

7. New round: Buyers receive new identification numbers. In the next round different product variants will be traded, i.e. the sellers' costs and the buyers' valuations will again be determined randomly. The valuations always lie between 10 and 20 ECUs and the costs between 0 and 10 ECUs, but the actual values change.

At the end of the experiment 10 of the 30 rounds are randomly selected for payment. Only these ten rounds are payoff relevant. Your average income of these 10 rounds will be converted into Euros. For sellers the exchange rate equals ECUs * 5 = €. For buyers the exchange rate is ECUs * 1 = €. In addition to their earnings from the selected 10 rounds sellers receive € 5 and buyers € 2.50 as a fixed payment. The earnings are paid out privately and in cash at the end of the experiment.

Chapter 3

Information and Communication in Posted Offer Markets: An Experiment*

3.1 Introduction

Market frictions may reduce buyer as well as seller surplus substantially. In markets with capacity constrained sellers, miscoordination is a friction with severe welfare consequences: when demand at a specific seller exceeds supply, some of the queuing buyers cannot purchase.[1] At the same time, other sellers are unable to find a trading partner, because potentially interested buyers have decided to approach an overcrowded seller. Thus, too few transactions take place and overall welfare suffers. Another important and similarly common friction is limited information on the demand side: buyers who do not observe prices cannot base their purchase decisions on them. They behave as if they were extremely price-insensitive. Consequently, they purchase randomly, often paying more than necessary. While buyers suffer from this friction, sellers gain from it. Thus, both limited information and price insensitivity lessen competitive pressures among sellers and may induce them to raise their prices knowing that each uninformed buyer might accidentally purchase their expensive products.

[1] If the requested product is sold out, consumers would normally try to find another seller and thereby incur search costs. The underlying theoretical model and the experiment abstract from this possibility by assuming that the additional search costs are prohibitively large.

* I am grateful to Henrik Orzen, Jörg Oechssler, Tim Cason, Stefan Penczynski, Jürgen Huber and Sheen Levine as well as seminar audiences at the University of Mannheim, ZEW Mannheim, the ESA meeting in Jerusalem, Friedrich-Alexander-Universität Erlangen-Nürnberg and the HeiKaMaX Experimental Workshop in Mannheim for their valuable comments and suggestions.

Examples for markets with frictions both on the demand and on the supply side are ubiquitous. The supply side may be capacity constrained with respect to different limiting factors: in the case of doctors and hairdressers, for example, the constraint relates to time. In an hour, a doctor or a hairdresser can only serve a limited number of patients or clients. Restaurants are capacity constrained with respect to space, landlords only have a limited inventory, i.e. a limited number of apartments, and stores can only keep a limited number of items on stock. These supply side frictions are often accompanied by coordination frictions on the demand side: patients need to decide which general practitioner to visit without knowing how many other patients are already sitting in his waiting room. When deciding at which restaurant to eat, people cannot be sure if it is not overcrowded. Potential tenants do not know how many other tenants have decided to view an empty apartment before they arrive at the appointment. In addition to the coordination problem, limited information constitutes another friction on the demand side. Patients may not know which share of the treatment is covered by their insurance, restaurant visitors who did not check the menu are unaware of the price level of the restaurant and some potential tenants have trouble estimating the total rent including all additional costs. In markets with capacity constrained sellers and coordination problems among buyers, constellations with too little trade may arise. I investigate these markets more closely in a controlled laboratory setting in order to find out how changes related to information and coordination affect buyer welfare.

Since enhancing information or facilitating coordination may attenuate the above mentioned frictions, the two measures are typically seen as welfare improving. In coordination games, communication has long been known to resolve the coordination problem (Cooper et al., 1989, 1992). As markets with capacity constrained sellers can be interpreted as coordination games, communication among buyers should reduce the coordination problem, increasing the welfare of both buyers and sellers. Yet in small markets with capacity constrained sellers, the consequences of better communication are not clear. When buyers can coordinate their purchase decisions, sellers are more likely to sell their goods than in the absence of coordination devices. In order to increase their chance of purchasing, buyers may be willing to seek higher priced products when they know that other buyers will purchase the cheaper products. Thus, coordination not only increases trade, but also induces sellers to raise their prices, because they can rely on being able to sell their goods. The overall effect of better coordination on buyer surplus therefore needs investigation: does communication improve buyer surplus through an increase in trade? Or does the negative welfare effect of coordination stemming from the price increase prevail?

A large body of research has come to the conclusion that providing more consumers with price information is consumer welfare improving (see e.g. Varian, 1980, Burdett and Judd, 1983 and Stahl, 1989). However, in small markets with capacity constrained sellers altering the number of informed buyers leads to two opposing effects: an additional informed buyer increases the overall price-sensitivity of buyers, because another buyer can identify and seek the cheapest product instead of purchasing randomly. Thus, more buyers react to price differences and fewer buyers (if any) decide randomly. Anticipating this increased price sensitivity, sellers will try to marginally undercut their opponent's price to attract potential buyers, because they can rely less on random purchases of uninformed buyers. Hence, providing an additional buyer with price information may cause some downward pressure on prices.

Lester (2011) shows that the opposite effect can also occur, i.e. more informed buyers can induce upward pressure on prices when sellers are capacity constrained: As each buyer knows that the other buyers can also observe the posted prices and thus can identify the cheapest product, competition for the chance to purchase cheap has intensified. Therefore, each buyer is aware that her likelihood to get to purchase the cheap product has decreased due to the change in price information. With capacity constrained sellers each buyer faces a trade-off when choosing a seller: if the buyer is able to purchase the good, she favors the cheapest product. The *likelihood* that she is able to purchase the cheap good, however, is lower when more buyers are informed, because another buyer might try to purchase from the cheapest seller as well. If she decides to buy from the more expensive seller instead, her consumer surplus from purchasing is lower, but the expected likelihood to be able to purchase is higher. Consequently, each informed buyer becomes more willing to purchase a more expensive good aiming to increase her likelihood to trade. Anticipating this decrease in the individual price sensitivity, sellers are induced to raise their prices. As these two effects – the increase in competition among sellers and the decrease in individual price sensitivity – point into opposite directions, the net effect is a priori not clear. Lester (2011) shows that the rather counterintuitive effect of more information to raise prices prevails in markets with two sellers and three buyers of whom one may or may not be informed about prices: with three informed buyers, prices are higher than in markets with two informed buyers and one uninformed buyer ("Lester's paradox").

In this paper I seek to analyze experimentally which of the two effects dominates. Do prices really increase when an additional buyer is informed? In addition, I investigate whether coordination through communication has an influence on the pricing and purchase decisions, focusing on the effect of coordination on buyer surplus. Will buyers suffer from better buyer coordination? Or will the positive welfare effect from an increase in trade prevail? Furthermore, I ask whether the impact of communication on buyer welfare depends on the informational setting. I also investigate the use of the option to communicate more closely. Will buyers avoid communication due to its increasing effect on prices? Besides, I seek to find out which patterns in communication lead to coordination.

In my experiment I consider a market with two capacity constrained sellers facing three buyers of whom one does or does not lack price information. All participants know whether an uninformed buyer is present in their market. Each seller offers exactly one unit of an indivisible good at a price that he determines at the same time as the other seller. The goods are homogeneous such that the offers can only differ in their prices. Once both sellers have set their prices, they are revealed to the informed buyers. Each buyer can purchase exactly one unit of the good and makes her purchase decision without observing the other buyers' choices. In treatments with communication I introduce an additional stage which gives buyers the possibility to coordinate if they wish. Before they make their actual purchase decision, buyers can inform the other buyers which seller they intend to choose, but they also have the option to refrain from communicating their purchase plans. Communication is cheap talk, i.e. buyers do not need to commit to their stated intentions. Buyers communicate their purchase plans sequentially before they choose from whom to purchase. If they communicate truthfully, coordination in the market should be improved. In case a seller finds more than one customer, the trading buyer is determined randomly while the other buyer(s) who chose the same seller cannot purchase. The income of a trading buyer is equal to her willingness to pay minus the price of the purchased good. Buyers who cannot purchase have no income. By the same reasoning, if a seller does not find a buyer, he cannot sell his good and has no income. In this setting I will investigate the effects of price information and buyer communication on the price level and consumer welfare.

So far, the literature has not investigated the effects of communication in small markets with capacity constrained sellers, but there are two experimental papers seeking to find evidence of "Lester's paradox" which come to different conclusions: Whereas Helland et al. (2017) (henceforth HMP) find higher prices in markets with three informed buyers than in markets with two informed buyers and one uninformed buyer and therefore agree with Lester's prediction, prices in the experiment of Anbarci and Feltovich (2017) (henceforth AF) are higher when one uninformed buyer is present contradicting Lester's model. Given the current state of the literature, it is not clear which finding is more robust. Furthermore, the effects of improved coordination due to communication need examination.

In my experiment subjects play the market game for 35 rounds followed by a short lottery selection task (see Eckel and Grossman, 2008) to elicit risk preferences. I use a 2x2 design with four between-subject treatments altering communication and the number of informed buyers. I find that communication has stronger effects on prices than information. In the treatments with communication I observe higher prices when all buyers are informed compared to the situation when one buyer lacks price information. Without communication prices are lower when all buyers are informed which contradicts Lester's prediction. Although communication increases trade, it leads to a *decrease* in buyer surplus because sellers set substantially higher prices. Thus, giving buyers the possibility to communicate and exchange their purchase plans is actually harmful to them. This finding is particularly interesting as communication is generally seen as welfare enhancing. Despite the harmful effect of the achieved coordination on their surplus, buyers do not stop communicating. Coordination is reached by honest communication and it is more likely to occur when the posted prices do not differ much. In general, communication influences the market outcome at least as much as information in posted offer markets with capacity constrained sellers.

The remainder of this paper is organized as follows: In section 2, I will summarize the related theoretical and experimental literature. Then, I will explain my experimental design and in section 4, I state my main research questions. I will present my results in section 5, starting with summary statistics and continuing with both a parametric and a non-parametric data analysis. Then, I investigate coordination more closely. The last section concludes.

3.2 Related Literature

This paper relates to various strands of the literature. One theoretical strand, which only treats the information friction, deals with buyers who may lack price information and face sellers without capacity constraints (see e.g. Varian, 1980, Burdett and Judd, 1983 and Stahl, 1989). Thus, it only treats the information friction. A common feature of many of these models is that in equilibrium, sellers randomize over prices. When the share of informed buyers increases, more buyers make informed purchases and fewer buyers decide randomly. Additional price information increases the average price sensitivity in the market and sellers stop relying on random purchases by uninformed buyers. This increase in price sensitivity leads to a rise in competition among sellers and induces them to reduce their prices which increases consumer welfare. Experiments that have investigated sellers' pricing decisions when buyers lack price information (e.g. Cason and Datta, 2006 and Morgan et al., 2006) find strong evidence of the theoretical predictions: prices decrease when more informed buyers are present in a market.

Another strand of the literature focuses on the coordination friction, considering capacity constrained sellers who interact with fully informed buyers (Burdett et al., 2001). In their model, buyers trade off the price with the probability of obtaining the good: as sellers are capacity constrained, situations in which sellers cannot serve all interested buyers may occur. Whenever queues form at one seller, too little trade occurs in the entire market, because some of the queuing buyers and sellers, who have not found any buyers, cannot trade. Anticipating not being served when choosing the cheapest seller, buyers also take more expensive sellers into consideration. Each buyer has to balance buying at a lower price with a lower probability against buying at a more expensive price with a higher probability. The price elasticity is therefore lower and the market price is higher when capacity constraints and coordination frictions are present compared to markets without capacity constrained sellers.

Cason and Noussair (2007) study the problem of simultaneously occurring excess supply and demand in an experiment by taking Burdett et al. (2001)'s model to the laboratory. They observe that price setting behavior in the laboratory generally resembles the predicted behavior, even though prices sometimes exceed the predictions and do not always converge to a common price. Especially in the treatment with two sellers facing three buyers, buyers seem to respond too little to price differences. The authors explain this finding with each buyer's rather low equilibrium probability of making a purchase: As buyers perceive the event of not being able to purchase as salient, they experience the competition for the chance to purchase as very fierce. Accepting higher prices to increase their purchasing probability, they induce sellers to set higher prices than predicted by the model. Apart from the price level, Burdett et al. (2001)'s model predicted the experimental data rather well.

Lester (2011) combines the above mentioned strands of theoretical literature by incorporating both information frictions and coordination frictions in a theoretical model. He investigates the pricing decisions of capacity constrained sellers facing buyers who do or do not observe the posted prices. Price-responsiveness varies both with the number of buyers relative to the number of sellers and with the proportion of informed buyers. In a setting with two sellers and three buyers of whom at least two have price information, the market price increases when exchanging the uninformed buyer with an informed buyer. This counter-intuitive result stems from the fact that informed buyers face the trade-off of getting a higher consumer surplus when they choose a lower priced seller versus buying from a higher-priced seller with a higher probability. When more buyers observe the prices and thus choose strategically instead of randomly, each buyer becomes less price sensitive, because she knows that another informed buyer might try to purchase from the cheaper seller. In this particular setting, the individual decrease in price sensitivity exceeds the increase in average price sensitivity: even though more buyers observe the prices and can base their purchase decisions on prices, competition among buyers for the chance to purchase induces each buyer to accept higher prices. Anticipating buyers' behavior, sellers raise their prices when more buyers are informed. I try to find experimental evidence for this prediction and introduce communication to the setting. Furthermore, I analyze whether buyer communication weakens or even reverses the effect

At least two papers have also investigated Lester's paradox experimentally and have found opposite effects: Anbarci and Feltovich (2017) study the effects of increasing the share of informed buyers on prices. They test whether prices in markets with one instead of zero uninformed buyers are actually lower. In their experiment, they consider markets with two sellers facing two or three buyers of whom at most one is uninformed. The theoretical prediction for the market with *two* buyers is that moving from zero to one uninformed buyer will *increase* prices. In a market with *three* buyers, however, moving from zero to one uninformed buyer will *decrease* prices. AF alter the number of uninformed buyers in a within-subject setting to remove confounds due to unobserved individual specific attributes. Hence, each subject passes through 20 rounds of a market game without uninformed buyers and 20 rounds of a market game with one uninformed buyer. They find that increasing the number of uninformed buyers leads to higher prices and profits in both markets with two and with three buyers, the latter finding contradicting Lester's prediction. The authors suggest that fairness is a possible explanation for their findings: when asked to report a "fair" price, both buyers and sellers state higher fair prices when one instead of zero uninformed buyers is present. Considering a higher price fair, sellers post higher prices and buyers become less price sensitive. AF conclude that it is the fairness views of the subjects, not Lester's paradox which drives the prices in their experiment.

Helland et al. (2017) also look for evidence for Lester's paradox in experimental posted offer markets. In contrast to AF, they use a between-subject design such that each subject is either part of a treatment with one uninformed buyer or part of a treatment without uninformed buyers. The market game is played for 50 rounds and computerized subjects take the role of uninformed buyers deciding randomly from which seller to purchase. Despite these rather small changes in the experimental design, HMP find evidence for Lester's paradox in the data. Although the effect is small, increasing the number of informed buyers leads to a significant rise in transaction prices. Following their example, I also use a between-subject design, but in my experiment human subjects take the role of uninformed buyers.

Apart from the experimental literature on posted offer markets, there is a vast experimental literature on the effects of communication on coordination. Van Huyck et al. (1993) find evidence that tacit pre-play communication fosters coordination. Blume and Ortmann (2007) show that cheap talk facilitates coordination and disentangle the different mechanisms through which cheap talk is effective, e.g. self-commitment through messages, negotiations of equilibria and secret handshakes. Departing from the literature on the entirely beneficial effects of communication, Cason et al. (2012) raise the issue that communication may have both negative and positive effects. In their setting communication can enhance or damage efficiency through inter- and intra-group communication.[2] Their research is therefore similar to this paper as it includes a more differentiated view on the effects of communication.

In this paper I investigate how the presence of more informed buyers in a market affects sellers' pricing decisions. I analyze the impact of communication on price setting and trade, focusing especially on buyer welfare. Furthermore, I examine the joint effect of communication and information on prices.

3.3 Experimental Design

The experiment was programmed in zTree (Fischbacher, 2007) and run at the laboratories of the University of Mannheim (mLab) and the University of Heidelberg (IWA-Lab).[3] Subjects were recruited through ORSEE (Greiner, 2004) in Mannheim and hroot (Bock et al., 2014) in Heidelberg. At both universities, the subject pools mainly consist of undergraduate students from various fields. In total, 400 subjects participated in 26 sessions. Large sessions were split into two matching groups to obtain two independent observations from one session. With a duration of 75 minutes, sessions with communication took slightly longer than sessions without communication. On average, subjects

[2] Intra-group communication improves coordination within the communicating group, but induces negative externalities on the competing group. The reason for this welfare loss is that communication enables group members to coordinate on higher efforts which increases the expected payoff within the group, but at the same time it reduces the expected payoffs of the competing group.

[3] There are no significant differences between data from Mannheim and Heidelberg.

earned €13.35. To create public knowledge of the instructions[4], they were read aloud to all participating subjects at the beginning of each session. Afterwards, subjects were given the possibility to ask questions.

The market game is played for 35 rounds. In each round subjects are randomly matched into markets of five participants. This design is intended to approximate one-shot situations while allowing for learning. Each market consists of two sellers and three buyers and subjects keep their roles as buyers and sellers throughout the whole session. Individuals are not given any labels, hence they cannot identify the other individuals over the rounds. This anonymity and the random matching protocol ensure that subjects are unlikely to collude.

Each round starts with sellers simultaneously setting their prices. Prices have to be chosen from the interval of [0, 20] ECUs and can be entered with up to two decimal places. After both sellers in a market have set their prices, buyers make their purchase decisions. At the end of a round all players are informed about their outcome in the respective round: if a seller has found at least one buyer, he earns the price he has demanded. Furthermore, he is informed about his rival's price and the number of buyers he was able to attract. In case he has not attracted a buyer, he does not earn anything in that round. If a buyer is the only buyer at a seller, she earns 20 ECUs minus the price of the selected seller. In case that more than one buyer wants to purchase from the same seller, one of the queuing buyers is selected at random. The buyers who have not been selected to purchase do not earn anything in the current round. Buyers are informed about the number of rivals at their selected seller, the prices in the market (which is new information for the uninformed buyers), whether they were able to acquire the good and which resulting payoff they have earned. Afterwards, a new round begins.

The experiment consists of four treatments. I use a 2x2 design varying the possibility to communicate and the information buyers can observe. Each subject took part in only one of the treatments. In the treatments without communication, buyers directly make their purchase decisions, while in the treatments with communication, buyers can announce to the other buyers from which seller they intend to purchase after observing (or not observing) the prices. These announcements are non-binding and are made sequentially to facilitate coordination. If buyers communicated simultaneously instead, more miscoordination would most likely arise, because buyers could not react to the announcements of other buyers when deciding which message to send. Communication

[4] See the appendix for a translation of the instructions.

is very stylized such that buyers cannot chat with each other, but instead they are given three pre-formulated options to choose from: they can communicate (i) that they are planning to purchase from seller A, (ii) that they are planning to purchase from seller B or (iii) that they prefer not to communicate their intentions. One of the three buyers is chosen randomly to start the communication sequence, i.e. she communicates her purchase plans first. Which buyer is determined to communicate at what position in the sequence varies from round to round. The second buyer in the communication sequence can observe the message sent by the first buyer when deciding which message to send herself. The third buyer observes both messages before she can announce her own plan. In the treatments with one uninformed buyer, the uninformed buyer is always last in the communication sequence. As soon as the third buyer has announced her plan, the three buyers in one market observe the sequence of the communicated messages and make their purchase decisions simultaneously.

In the treatment with partial information one of the buyers is not informed about the prices, while in the full information treatments all buyers can observe the prices before making their purchase decisions. The uninformed buyer has to select a seller without knowing how much she will need to pay. Which subject takes the role of the uninformed buyer is determined randomly in each round, such that in treatments with partial information, every buyer is sometimes uninformed. The informed buyers can observe both prices and can decide if they want to purchase from the cheaper seller or rather from the more expensive seller hoping they will be alone in the queue at the seller.

After subjects have played the market game for 35 rounds, they receive new instructions before a lottery selection task starts on their screens (Eckel and Grossman, 2008). Lotteries vary in their degree of riskiness and in their expected value with the safe lottery yielding €4 no matter which outcome occurs and the most risky of the five lotteries resulting either in a payoff of €12 or in a payoff of €0 with equal probabilities (see Table 3.1). Subjects choose their preferred lottery, thereby revealing some information about their risk preferences to the experimenter.

At the end of the experiment, ten payoff-relevant rounds are randomly selected from the 35 rounds in the first part of the experiment. Subjects are shown a list of their earnings in all periods to verify that their payoff has been computed correctly. The average of the earnings in Experimental Currency Units (ECUs) from the selected rounds is taken and converted into Euros using different exchange rates for sellers and buyers which are common knowledge. For sellers, an ECU is multiplied by 2/3, whereas for buyers, an ECU is multiplied by 8/3. The difference in exchange rates ensures

Table 3.1: Lottery selection task

	1-50	51-100
Lottery 1	€4	€4
Lottery 2	€6	€3
Lottery 3	€8	€2
Lottery 4	€10	€1
Lottery 5	€12	€0

comparable earnings for sellers and buyers. After subjects have been informed about their total earnings in the experiment, a short questionnaire about demographics and strategies starts on their screens. Before leaving the laboratory, subjects receive their earnings privately and in cash.

3.4 Research Questions

Q 1. Are prices higher in treatments with full price information compared to treatments with partial price information?

Lester's (2011) model[5] predicts that in markets with two sellers and three buyers, sellers set higher prices when all buyers have price information compared to the case when one of the three buyers cannot observe the prices and has to decide randomly. While indifferent in equilibrium, informed buyers face a trade-off between the higher individual consumer surplus they receive when purchasing from a cheap seller and the possibly higher probability of purchasing when they visit the more expensive seller. Hence, having three instead of two informed buyers has two effects on the price sensitivity in the market: on the one hand, it increases price sensitivity because now all buyers observe the prices and base their decisions on them. On the other hand, it may also have a decreasing effect on the average price sensitivity because now each informed buyer is more likely to buy from the expensive sellers since she knows she is competing with other informed buyers for the chance to purchase. The experiment is designed to find out which of the two effects prevails.

[5] A brief description of the theoretical background can be found in the appendix.

Q 2. Are prices higher in treatments with communication than in treatments without communication?

When buyers communicate before purchasing, they can better coordinate their decisions. Consequently, the situation of all buyers trying to purchase from one seller while the other seller is left without customers should occur less often. *Without* communication, buyers use mixed strategies in the model visiting each seller with a probability that depends on the respective prices. If communication induces perfect coordination, the informed buyers' purchase strategies may become asymmetric: one informed buyer visits one seller with certainty, whereas the other informed buyer chooses the other seller with a probability of one. The uninformed buyer will randomize between sellers and the third informed buyer will purchase from the cheaper of the two sellers. With *perfect* coordination, sellers will always find at least one buyer. As they own only one unit of their product, they do not seek to attract more than one buyer. Starting with identical prices below ECU 19.99, each seller has an incentive to set a slightly higher price than his competitor as long as he will not lose his informed buyer. Knowing that another buyer plans to purchase from the cheaper seller, the buyer who planned to purchase from the more expensive seller can either switch to the cheaper seller or remain at the slightly more expensive seller. When she purchases from the cheaper seller, she receives a higher payoff, yet the probability to be able to purchase is lower. Thus for small price differences, she will buy from the slightly more expensive seller. Since both sellers anticipate this reasoning, the other seller is also encouraged to raise his price. Hence, prices may approach the maximum price of ECU 20 if communication induces perfect coordination. At this price, buyers are indifferent between coordinating and not coordinating because both strategies yield a payoff of zero. If they choose to coordinate, sellers earn their maximum payoff. At a price of ECU 19.99, buyers have a strict incentive to coordinate, because not buying yields a strictly lower payoff. As switching to the other seller weakly decreases their probability to make the purchase, buyers have an incentive to follow their announced purchase plan. Sellers earn (almost) their maximum payoff. Coordination thus softens competition among sellers and encourages them to raise their prices. In theory, sellers will charge the maximum price when they can rely on finding a buyer with certainty. The experimental price level thus depends on the level of coordination buyers reach by communicating.

Q 3. Do buyers decide not to communicate?

Assuming that communication facilitates coordination, it is less likely that all buyers decide to buy from the same seller when communication is possible. However, this improved coordination creates an incentive for sellers to increase their prices. In each round, buyers decide whether to communicate when prices have already been set. *Within* a round, they always have an incentive to achieve coordination as long as prices are (marginally) below their maximum willingness to pay since coordination increases the probability with which they can purchase. *Across* rounds, coordination has an adverse effect on prices. Realizing this, sophisticated buyers may deliberately decide not to communicate their purchase plans in order to *avoid* good coordination. If all buyers in a market decide to resist their incentive to coordinate for given prices, prices could decrease, but this decrease comes at a cost for buyers: without communication, miscoordination is more likely which decreases the likelihood of trade. Hence, I am curious to find out if buyers use the option not to communicate differently throughout the rounds of the experiment. Another possibility of using communication to influence the price level might be to coordinate on the cheaper seller. By not purchasing from the more expensive seller buyers can punish him and induce him to set a lower price in the following round. However, this strategy also requires a certain degree of sophistication among buyers who need to sacrifice their current payoff in order to potentially earn more in future rounds. The fact that new markets are formed in each round complicates this punishment strategy, yet I seek to find out if there is evidence for this behavior in the data.

Q 4. Do buyers communicate truthfully?

An improvement in coordination leading to the above mentioned consequences is only possible if buyers communicate truthfully. However, it is also possible that sophisticated buyers deliberately misinform the other buyers in their market in order to make communication less effective and thus avoid a price increase.

Q 5. Does communication enhance buyer welfare?

The effect of communication on buyer welfare is ambiguous: as communication facilitates coordination, I expect more trade to occur in treatments with communication. Through increasing trade, communication has a positive impact on consumer welfare. However, it may also reduce consumer welfare if prices are higher. I seek to find out which of these effects is stronger or if they neutralize each other. Focusing on sellers, I expect communication to enhance seller surplus, because better coordination leaves sellers less frequently without trade and induces them to set higher prices. Therefore, profits and total welfare will be higher when communication is possible.

3.5 Results

Applying both parametric and non-parametric methods, I analyze the effects of information and communication on prices. Furthermore, I aim to understand how the option to communicate is used by buyers and I compare situations in which buyers manage to coordinate to situations in which all buyers try to purchase from the same seller. In particular, I ask what prices sellers post and how buyers communicate when buyer coordination is achieved or not achieved. Finally, I examine whether the option to communicate benefits buyers or if they fall prey to potential adverse effects of communication.

Unless explicitly stated otherwise, I will focus my analysis on rounds 21 to 30. It is likely that the changes in prices in the first rounds of a session are driven by learning effects. Excluding these rounds, I ensure that the treatment effects are not polluted by potentially erratic choices at the beginning of the experiment. Similarly, I do not include the last five rounds (rounds 31-35) to rule out possible end game effects.

3.5.1 How do communication and information influence prices?

Before analyzing the treatment effects in detail, I consider the overall effects of information and communication on the average price level. First, in the treatments with full information the mean posted price in rounds 21 to 30 is ECU 16.29 and in treatments with one uninformed buyer it is ECU 16.19. This difference is small and with a p-value of 0.489 in a two-sample Fisher-Pitman permutation test, it is also statistically insignificant. Second, in the treatments with communication the average price is ECU 16.67 compared to a price of ECU 15.63 when buyers cannot communicate. This difference is statistically significant ($p = 0.010$). Thus, giving buyers the option to communicate induces sellers to increase their prices.[6]

Result 1: Communication increases the average posted price.

Next, I disaggregate the data and compare the observed and predicted prices for each of the four treatments (see Table 3.2). I find that observed prices exceed predicted prices for settings without communication, both when considering all rounds together and when focusing only on round 21 to round 30. Table 3.2 also shows results from previous experiments that tested Lester's paradox. Both AF and HMP observe lower prices than I do with HMP's prices coming very close to predicted price levels.

Cason and Noussair (2007) provide a potential explanation for high prices in experiments with capacity constrained sellers: due to the perceived salience of no trade, buyers fear not being able to purchase and competition among buyers for the possibility to

[6] If communication induced perfect coordination, one would expect prices close to ECU 20, as discussed above. Whether coordination was indeed perfect will be answered later in this section. However, sellers' fairness concerns might prevent them from setting prices close to ECU 20.

Table 3.2: Overview of predicted and observed prices

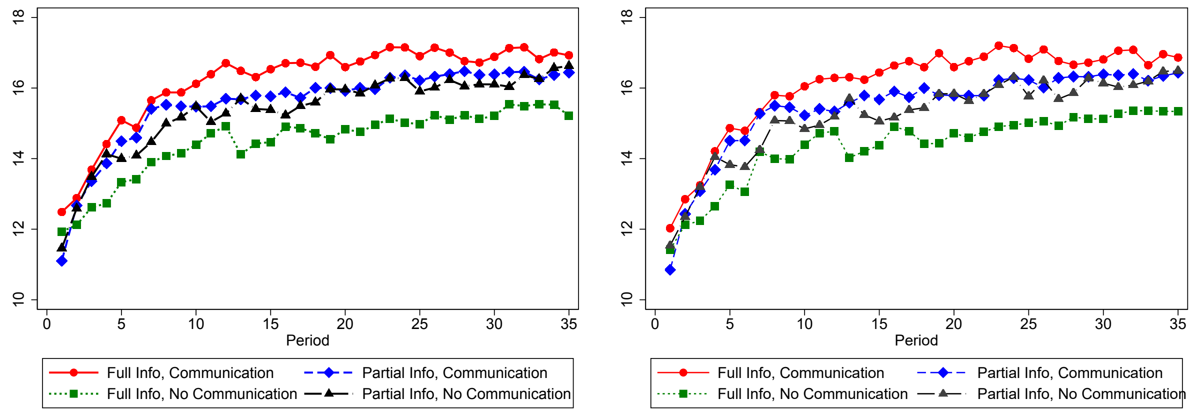
Setting	Full Information	Partial Information	Study
predicted	14.55	13.33	Lester
all rounds			
Communication	16.21	15.56	*this paper*
No Communication	14.49	15.36	*this paper*
	14.38	13.78	Helland et al.
	11.86	13.19	Anbarci & Feltovich
rounds 21-30			
Communication	16.94	16.28	*this paper*
No Communication	15.07	16.09	*this paper*
	14.70	14.07	Helland et al.
last 5 rounds	12.98	14.19	Anbarci & Feltovich

purchase increases. By accepting very high prices they induce sellers to raise their prices. Risk aversion of buyers could also explain the high price level, but Cason and Noussair doubt this explanation. Risk averse buyers accept higher prices than risk neutral buyers, but if sellers were also risk averse, they would set lower prices which cannot be found in the data. Thus, with equally risk averse buyers and sellers, the net effect of risk aversion on the price level is unclear. My own measurements of risk preferences do not differ systematically between sellers and buyers.

Furthermore, Table 3.2 shows that the average price level depends on the treatment: in the treatment with communication and full buyer information, prices are *higher* than in the treatment with communication and partial buyer information. Considering the treatments without communication, the prices in the treatment with full information are *lower* compared to the treatment with one uninformed buyer which seems to contradict Lester's paradox. Communication increases prices no matter which information setting is present. Thus, giving buyers the possibility to coordinate may harm them. In addition, it seems as though communication alters the effect of more information.[7]

As can be seen from Figure 3.1, sellers set rather modest prices at the beginning of a session splitting the welfare almost equally between themselves and the buyers. Within the first rounds they raise their prices significantly, but after about ten rounds, the rate of increase slows down. Posted prices (left panel) are the prices sellers post at the beginning of a round hoping to attract at least one buyer. In contrast, transaction prices (right panel) are the prices at which trade has actually taken place. Transaction

[7] This effect will be studied more deeply in the parametric analysis.



(a) Posted prices **(b)** Transaction prices

Figure 3.1: Time series of prices

prices evolve in a similar pattern as posted prices: at the beginning of a session, trade takes place at moderate prices. Over the rounds, sellers set higher prices and buyers accept these higher prices. Uninformed buyers do not know in advance at which prices they purchase, hence they are never discouraged from high prices. I observe increasing transaction prices in all treatments, but the increase levels out in the later rounds of the sessions. Since transaction prices demonstrate the same patterns as posted prices, I will use posted prices in the analysis.

Non-Parametric Tests

In order to compare the price levels across the treatments, I use non-parametric two-sample Fisher-Pitman permutation tests. As shown in Table 3.3, I find that the effect of information on prices depends on the presence of communication: with communication, prices are higher when all buyers are informed, whereas without communication, prices are lower with full information which contradicts Lester's paradox. While the latter effect is marginally significant, the former effect is not significant at any conventional level. Note that with at least seven matching groups per treatment, my number of statistically independent observations exceeds the number of observations in previous experiments. In the treatments with one uninformed buyer, communication does not seem to change prices, but when all three buyers are informed about prices, communication increases the prices significantly.

Table 3.3: Posted prices: Results from Fisher-Pitman permutation tests

Setting	Full Information	Partial Information	p-value
Communication	16.94	16.28	*0.133*
No communication	15.07	16.09	*0.094*
p-value	*<0.001*	*0.340*	

Regression Analysis

In order to control for risk aversion, experience and gender, I perform a regression analysis clustering the standard errors at the level of matching groups (see Table 3.4). I regress the sellers' posted prices on dummy variables specifying the treatment conditions, on the round number to include experience and on the squared round number to account for non-linear effects. The dummy variable "Information" is equal to 1 for the treatments with three informed buyers. Similarly, the "Communication" dummy is 1 when buyers had the possibility to communicate their purchase intentions. I interact these two dummy variables in the "Info * Communication" dummy. Furthermore, I use the chosen lottery as a proxy for the risk preferences of a seller: the higher the number of the chosen lottery, the more risk loving is the seller. As I want to analyze the impact of the sellers' and the buyers' gender separately, I build two gender variables to include gender effects. First, I control for the seller's gender using a dummy variable which is equal to 1 when the seller is male. Second, I include a variable that represents the gender composition of buyers in a market. It counts the number of male buyers in a market.

Regressing the posted price on the dummy for "Information", the dummy representing "Communication" and their interaction while controlling for experience by including the round number and the squared round number I find that the coefficient of the information dummy is negative and insignificant, communication has a positive, but insignificant impact and the coefficient of the interaction term is positive and significant. Hence, information decreases prices (but not significantly so) in treatments without communication which is at odds with Lester's paradox. No significant effect of communication can be found for treatments with partial information[8], but the interaction term of communication and information has a significant effect on prices: in the treatments with communication, prices are higher with full information than with partial information.[9] In the treatments with full information, communication increases the price level significantly. As depicted

[8] This effect was significant in the non-parametric analysis.

[9] The effect was not significant in the non-parametric analysis.

Table 3.4: Regression analysis

	(1) Posted Price	(2) Posted Price	(3) Posted Price	(4) Posted Price
Round Number	0.466**	0.466**	0.506**	0.506**
	(2.55)	(2.55)	(2.09)	(2.09)
(Round Number)^2	-0.009**	-0.009**	-0.010*	-0.010*
	(-2.45)	(-2.44)	(-2.04)	(-2.03)
Information	-1.016	-0.912	-0.571	-0.574
	(-1.45)	(-1.34)	(-0.90)	(-0.95)
Communication	0.189	0.290	0.459	0.457
	(0.24)	(0.38)	(0.58)	(0.59)
Info * Communication	1.678*	1.470*	1.764*	1.768*
	(1.93)	(1.72)	(1.99)	(2.05)
Seller's Lottery Choice		0.161*	0.183*	0.184*
		(1.78)	(1.82)	(1.79)
Male Seller			0.154	0.155
			(0.55)	(0.56)
Number of Male Buyers				0.008
				(0.04)
R^2	0.107	0.119	0.201	0.201
N	1600	1600	1120	1120

Output from OLS regressions, SE clustered on matching group level
t-statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in Figure 3.1, the round number has a positive, yet decreasing effect on prices. Adding the lottery choice as a control variable, I find that more risk-loving sellers choose significantly higher prices which is intuitive. There is no significant gender effect and the interaction term of communication and information stays significant when gender variables are considered in the model. Although the gender of sellers and risk preferences are positively correlated – male sellers are more likely to choose a risky lottery than female sellers – gender does not have a significant impact on the pricing decision.

Result 2: When all buyers observe the posted prices, communication has a positive effect on the price level. When one buyer lacks price information, the impact of communication on prices is insignificant.

Result 3: In the treatments without communication, no evidence for Lester's paradox could be found. If anything, it is the other way round.[10]

[10] This result relies on the non-parametric tests. Evidence for this effect was not significant in the regression analysis.

Result 4: In the treatments with communication, information has an increasing effect on prices.[11]

3.5.2 Do differences between sellers' prices vary across treatments?

Communication might facilitate coordination. When buyers coordinate, sellers do not need to undercut their opponent's price to attract the informed buyers. Thus, prices may be more similar when buyers can communicate. In the two treatments with three informed buyers, the average difference between the higher and the lower price is ECU 1.59 and the average price difference in the treatments with partial information is ECU 1.39. This difference is not statistically significant (p=0.225). When buyers have the option to communicate, prices are more similar than in treatments without communication (ECU 1.29 vs. 1.80). Here, the difference is marginally significant (p=0.076).

Result 5: Communication has a decreasing effect on the difference between the posted prices, while information does not affect the price difference.

Figure 3.2 shows that at the beginning of a session, the price differences are high across all treatments. When subjects gain experience, they set more similar prices. After about 15 rounds, it becomes obvious that the price difference without communication exceeds the price difference in treatments with communication. Towards the end of the experiment, however, the price difference in treatments without communication decreases towards the level in treatments with communication. This finding may result from the increase in experience among subjects: possibly, more experienced sellers are better at predicting their opponent's price and thus know what price they need to set to match or slightly undercut his price. Another possibility could be that tacit collusion plays a role. Even though sellers remain anonymous and new markets are formed after each round, sellers know that the number of all possible interaction partners is limited. As sellers are informed about their opponent's price after each round, they might learn to adjust their price.

When I consider each treatment separately, I observe that the mean price difference is highest in treatments without communication and full price information. Thus, the decreasing effect of communication on price differences is most pronounced when all buyers are informed about prices (see Table 3.5).

[11] This result relies on the regression analysis. Evidence for this effect was not significant in the non-parametric tests.

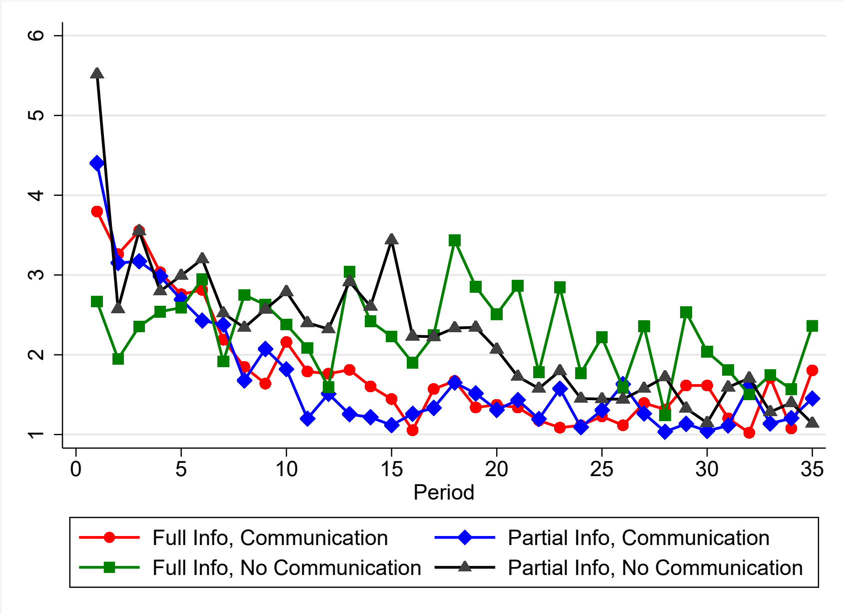


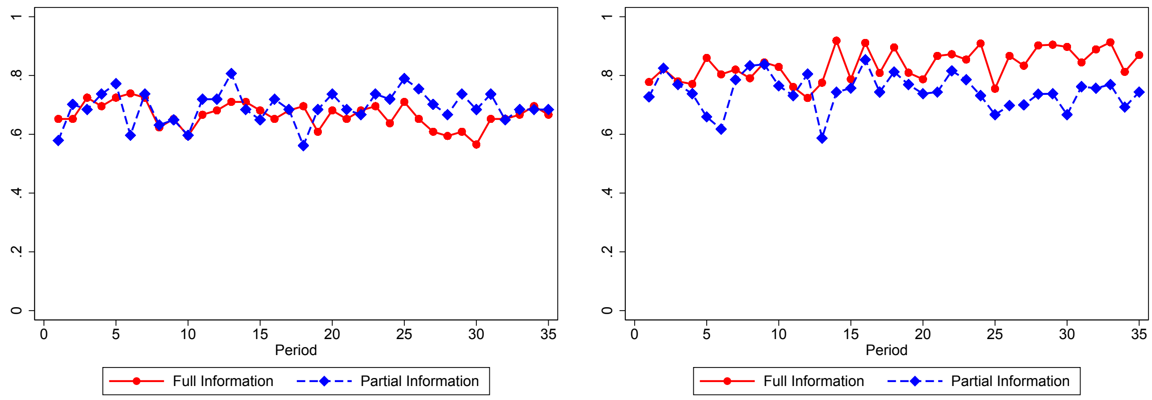
Figure 3.2: Time series of price differences

Table 3.5: Price differences: Results from Fisher-Pitman permutation tests

Setting	Full Information	Partial Information	p-value
Communication	1.30	1.27	*0.473*
No communication	2.12	1.52	*0.120*
p-value	*0.046*	*0.368*	

3.5.3 Communication

Communication may be used to improve coordination among buyers. The increase in the number of transactions comes at a cost for buyers as sellers may raise their prices. It might thus happen that more experienced buyers stop communicating their purchase intentions either by lying about their true plans or by choosing the message that they prefer not to communicate their plans. Yet buyers might also use communication as a tool to deliberately boycott expensive sellers and thus teach them to set lower prices. Figure 3.3 shows that the usage of the option to communicate does not change systematically throughout the rounds. We can therefore reject the hypothesis that more experienced buyers decide not to communicate. For given prices, buyers follow their incentive to coordinate. Besides, no significant treatment differences in the tendency to communicate can be observed (p = 0.136): With full information, on average 64% of the buyers use the option to communicate, whereas with partial information on average 71% communicate their intentions.



(a) Usage **(b)** Honesty

Figure 3.3: Communication

However, there are significant treatment differences concerning the honesty[12] of the communicated messages: with partial information, 73% of the messages are truthful, whereas with full information, 86% of the buyers purchase according to their stated intention (p = 0.015). Interestingly, this result is not driven by the uninformed buyers who communicate their intentions at the end of the communication sequence. Also the informed buyers who communicate first are more likely to deviate from their stated intentions when an uninformed buyer is present in their market. Yet the difference in honesty is only significant for the second informed buyer (p = 0.020).

Result 6: The majority of those who communicate contribute to better coordination by sending truthful messages. Purchase decisions are more consistent with the announced intentions in the full information treatment than in the partial information treatment.

Result 7: The communication behavior does not change with experience.

So far, I have only considered the usage of the options to communicate and the honesty of the messages, i.e. how often buyers stick to their announcements, over time. Now, I will focus on the *content* of the messages sent. As each buyer could choose between not stating her intentions (N), announcing that she will acquire the cheap product (C) and communicating that she will try to buy the expensive product (E), there are in total 27 possible strategy combinations between the three buyers. Figure 3.4 provides a closer look at the frequencies with which the strategies were used in the full information treatment (left panel) and in the partial information treatment (right

[12]A message is referred to as honest if the buyer purchased according to her announced plan. I do not differentiate between messages that stated the true intentions at the moment they were sent, but the buyer changed her mind after observing the other intentions and messages that were meant to mislead the other buyers.

panel). The length of each bar represents how often the first and the second buyer used a combined strategy. The first letter in the strategy descriptions denotes the message sent by the first buyer in the communication sequence and the second letter represents the message sent by the second buyer, who could observe the first buyer's message. For example, EN means that the first buyer announced the plan to purchase expensive whereas the second buyer did not communicate her plans. The third buyer could observe both messages when deciding which communication strategy to choose herself. The segments of the bars depict the percentages with which the third buyer employed each of the three messages.

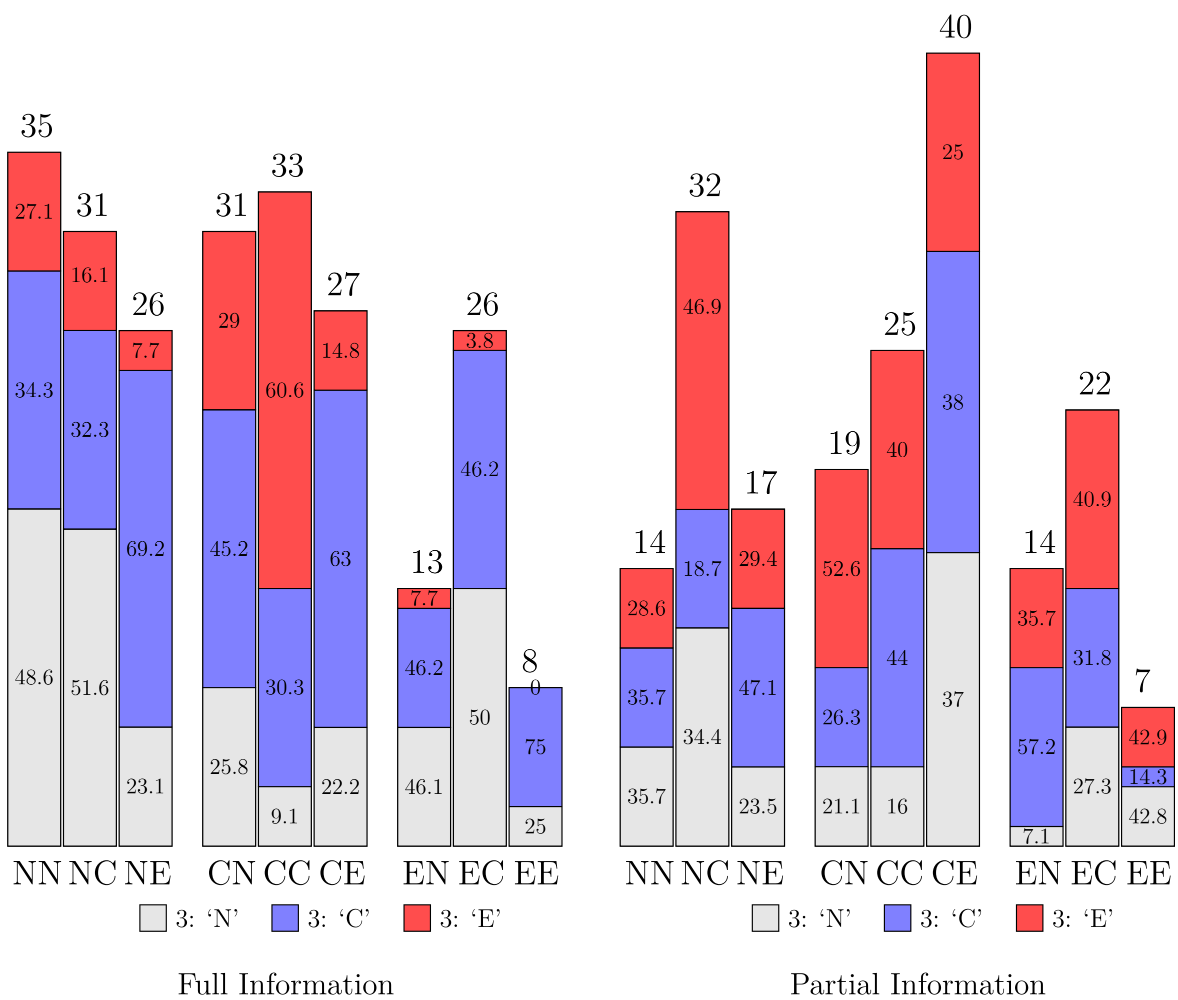


Figure 3.4: Messages of the 3^{rd} buyer related to messages of the first two buyers

It can be seen from the figure that the message combination which the first two buyers sent most often in the treatment with full information is NN followed by CC. In the treatment with partial information it is CE, i.e. the first buyer announced to purchase cheap; knowing this, the second buyer stated that she would purchase the expensive good. If the first and the second buyer communicated truthfully, the third buyer could expect to always compete with another buyer, but did not know which of the products was more expensive. In the full information treatment, the third buyer had to choose between receiving the expensive or the cheap product with a chance of 50%. If she trusted the messages of the other buyers, she ought to have preferred the cheaper product. And indeed, the most frequent answer of the third buyer to CE was C in the treatment with full information, with partial information, C and E occur with similar frequency because the third buyer cannot identify the cheaper seller. When at least one of the first buyers has communicated that she will purchase the expensive product, the third buyer avoids announcing that she will purchase expensive if she observes the prices. As can be seen from Figure 4, the first buyer tried to take advantage of her position in the communication sequence and intended to secure the cheap product by announcing to purchase cheap (CN, CC, CE) more often than announcing to purchase expensive (EN, EC, EE). This observation holds for both information conditions.

NC also constitutes a common strategy. Here, the third buyer mostly reacted by stating that she would purchase expensive in the partial information treatment[13] and she did not communicate at all in the full information treatment. It also happened frequently that the first and the second buyer competed for the cheap product (CC). In the full information treatment, the most common reaction of the third buyer was to announce that she would buy the expensive product. If communication was truthful, she would pay a higher price, but receive the product with certainty. Strategy combinations with two or three buyers announcing that they would purchase the expensive product were hardly ever used (ENE, NEE, EEE, EEN) in the treatment with full information.[14] In summary, the data suggests that buyers used communication to avoid miscoordination.

3.5.4 Coordination

In this section, I analyze the patterns in the prices set by sellers and the communication strategies used by buyers in order to better understand buyer coordination. Specifically, I want to find out whether buyers coordinate more often when sellers set lower or higher prices. Furthermore, I ask whether not only the price *level*, but rather the *difference* between the two posted prices in a market has an impact on coordination. I also take

[13] Of course, the third buyer did not know that her message meant that she was planning to purchase from the expensive seller. Instead, she rather avoided the seller with at least one client. Besides, she might have anticipated that the second buyer announced the plan to purchase cheap.

[14] The last letter denotes the strategy of the third buyer.

Setting	Full Information	Partial Information	p-value
Communication	0.77	0.77	**0.517**
No Communication	0.72	0.71	**0.704**
p-value	**0.332**	**0.265**	

Table 3.6: Coordination by treatment

a closer look at the communicated messages focusing on the buyer's position in the communication sequence and comparing the announced actions to the actual purchase decisions. Which messages did buyers send and from which seller did they purchase when coordination was achieved?

Table 3.6 summarizes the share of markets in which coordination has been achieved. The depicted shares are based on all rounds of the experiment. In a market with successful coordination, both sellers could sell their good and only one buyer was unable to purchase, whereas only one buyer and one seller could trade if miscoordination occurred. From the table it becomes clear that coordination was in fact easier to achieve when buyers could communicate. According to a two-sided two-sample Fisher-Pitman permutation test the null hypothesis of no difference in coordination between the treatments with and without communication may be rejected ($p = 0.098$). The results from comparing the treatments separately are insignificant.

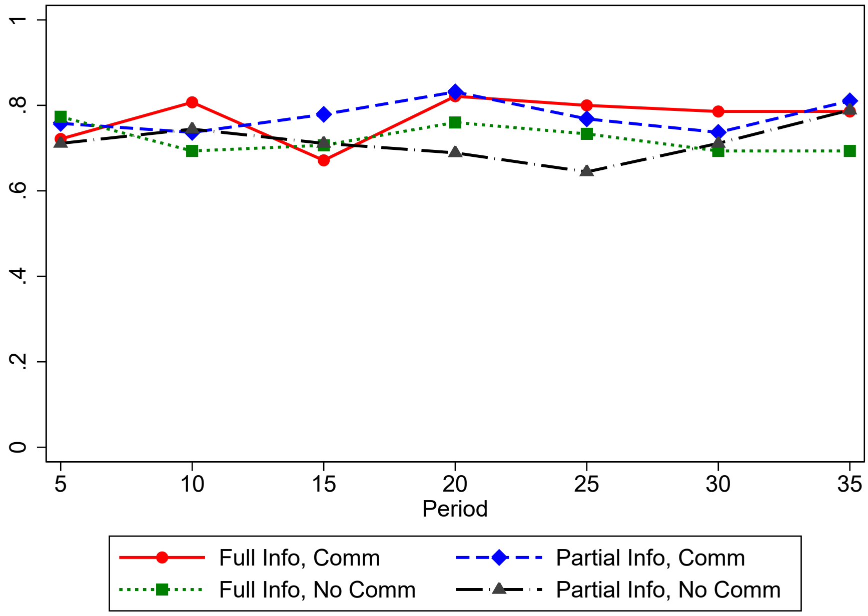
Figure 3.5 shows the average level of coordination over time for each treatment.[15] While the difference between treatments is small in the beginning of the experiment, after 15 rounds coordination is achieved more often in the treatments with communication. Yet towards the end of the experiment the share of markets with successful coordination increases in the treatment with partial information and without communication.

Does coordination depend on the average price level?

When the price level in a market is very high, buyers receive only a small payoff from purchasing. As the difference in payoffs between purchasing an expensive good and not being able to purchase (which leads to a payoff of 0) is small, buyers might care less for being able to purchase in comparison to situations with lower prices. Thus, they might put less effort into coordinating their purchase decisions.

[15]For the sake of better readability the graph depicts averages over five rounds.

Figure 3.5: Coordination over time



In order to find out if miscoordination is more likely to occur when prices are higher, I first compare the average price levels in situations with and without successful coordination. I find that the average price when coordination is successful slightly exceeds the average price when miscoordination occurs (16.27 vs. 16.16). Even though the difference is rather small, it is marginally significant at the 10%-level.[16] Analyzing the posted prices for each treatment separately, I observe no systematic differences between the posted prices with and without successful coordination (see Table 3.7).

Table 3.7: Average price levels with and without coordination

Setting	Coordination	Miscoordination
Full Information, Communication	16.99	16.74
Partial Information, Communication	16.22	16.48
Full Information, No Communication	14.98	15.31
Partial Information, No Communication	16.17	15.93

[16]This result stems from a one sample Fisher-Pitman permutation test. Considering only the first 20 rounds, the result stays marginally significant, but for the last 15 rounds the difference between posted prices with and without coordination is insignificant (p = 0.169).

Does coordination depend on the price difference?

Apart from the average price level, the difference between the cheaper and the more expensive posted price may also influence the buyers' ability to coordinate: when all buyers try to purchase from the same seller, each buyer receives the product with a probability of $\frac{1}{3}$. If one of the products is very cheap and the other comparably expensive, it can happen that the expected payoff of purchasing the cheap good exceeds the expected payoff of purchasing the expensive product. Thus, it may be optimal for buyers not to coordinate, but try to purchase the cheap product. It could also be possible that some subjects try to punish the expensive seller by purchasing from the cheaper seller, even though subjects are randomly matched into new markets in each round. These buyers might accept the lower expected probability of purchasing from the cheap seller in order to 'teach' the other seller to lower his price in future rounds. This behavior may also be reflected in a higher level of miscoordination when the price difference is high.

Aggregating over all treatments, the difference between the two posted prices is on average higher when miscoordination occurs (1.37 with coordination vs. 1.95 with miscoordination).[17] Hence, buyers find it harder to coordinate when one of the prices is substantially lower than the other posted price in their market. Focusing on the treatments with communication, the average difference is 1.15 when coordination was successful as opposed to 1.78 when coordination could not be achieved. In the treatments without communication, prices differ by 2.03 when coordination was not successful, while they differ by only 1.69 when coordination could be achieved. Although the price difference is higher in treatments without communication, it exhibits the same patterns as the price difference in treatments with communication. The price differences in situations with and without coordination can be found in Table 3.8 for each treatment separately.

Result 8: Buyers coordinate more often when the two posted prices do not differ much.

Does coordination depend on the buyers' communication strategies?

I seek to identify the communication strategies that facilitated coordination and compare them to those that hindered coordination. Figure 3.6 shows for each communication strategy in the treatment with full information by how many percentage points the corresponding mean achieved coordination deviates from the treatment mean coor-

[17] This difference is significant at the 1% level in a one sample Fisher-Pitman permutation test.

Table 3.8: Average price differences with and without coordination

Setting	Coordination	Miscoordination
Full Information, Communication	1.13	1.97
Partial Information, Communication	1.18	1.56
Full Information, No Communication	2.14	2.09
Partial Information, No Communication	1.30	1.98

dination of 79%. The strategy combinations that always resulted in coordination were NEE, CEN, CEE, ECN, ENE and ECE, yet some of them were hardly ever used (e.g. ENE). The strategies sticking out due to low levels of coordination are $NNC, NCN, NCC, CCN, ENN, ENC$ and EEN: When they were used, buyers could achieve coordination in less than 66% of the situations. As can be seen from Figure 3.6, strategies involving no communication from the first buyer are on average less successful in achieving coordination than strategies in which the first buyer states her intentions. However, if all buyers announced that they would purchase cheap (CCC) or only the first two buyers announced that they would purchase cheap (CCN) and the last buyer did not communicate, the majority of buyers – independent of their position in the sequence – actually tried to purchase cheap causing high levels of miscoordination. Yet buyers may have caused this apparent miscoordination on purpose in order to teach expensive sellers to set lower prices.

Since the last buyer in the communication sequence reacts to the already communicated messages, I focus on the first two buyers in the sequence. Figure 3.7 visualizes their behavior: the length of the three bars represents the frequency with which all first buyers in rounds 21 to 30 have chosen not to communicate, announced that they intent to purchase cheap or claimed that they would buy expensive. The segments of the bars show the second buyers' answers to the corresponding first-buyer messages. The pie charts below each bar segment depict how often coordination resulted from the respective strategy combination.

The first buyers preferred not to state their intentions in 36.9% of the situations, they declared that they will purchase cheap (41.7%) or they announced that they plan to purchase the expensive product (21.4%). If the first buyer has not made use of the communication option, the second buyer mostly answers that she will purchase cheap (40.7%) or does not communicate her plans either (31.6%). As expected, the

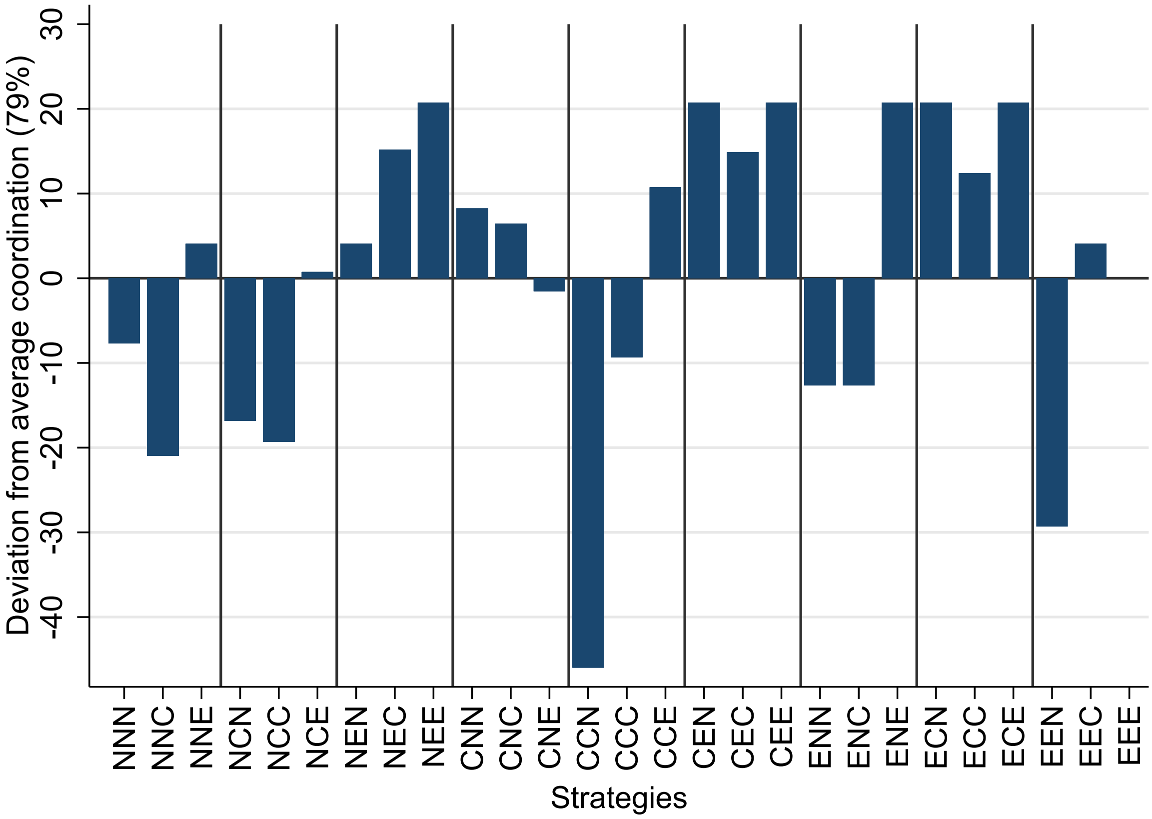


Figure 3.6: Coordination in relation to communicated messages

most common answer to the first buyer announcing to purchase cheap is the second buyer stating that she will purchase the expensive good. Similarly, when the first buyer plans to purchase the expensive good, the second buyer mostly answers that she will buy the cheap product. With coordination following in 86.6% of the situations, the first buyer saying 'cheap' and the second buyer answering 'expensive' is the most successful communication strategy, followed by 'cheap, no communication' (84%) and 'expensive, cheap' (83.3%). When buyers communicated 'cheap, expensive' or 'expensive, cheap', they followed their announced intentions in 70% of the situations. Thus, the majority of their announcements were truthful. The least successful strategy in achieving coordination is 'no communication, no communication' with coordination occurring in only 67.4% of the cases when both the first and the second buyer did not use the option to communicate.

Result 9: Truthful communication facilitates coordination.

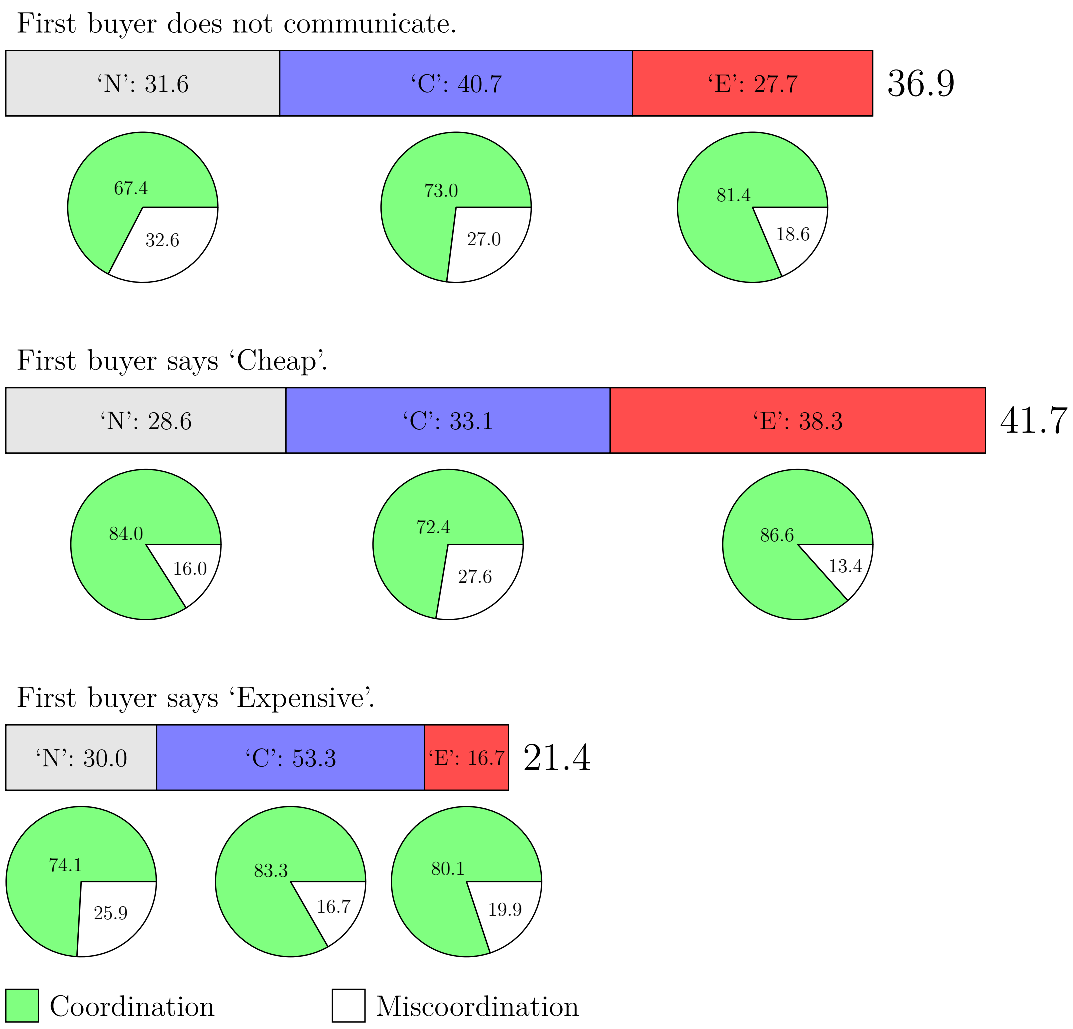


Figure 3.7: Coordination by communication strategies

3.5.5 Welfare Analysis

I analyze the impact of communication on welfare in Table 3.9. Since communication facilitates coordination, trade increases by six percentage points. As sellers are more likely to find a buyer in treatments with communication, they set higher prices. While the effect of communication on *coordination* has positive welfare consequences for both sellers and buyers, the effect of communication on *prices* shifts welfare from buyers to sellers. As expected, sellers benefit from buyer communication: miscoordination occurs less frequently and buyers accept higher prices. For buyers, the negative effect of communication prevails: they suffer more from the increase in prices than they benefit from improved coordination.

Result 10: Communication has a negative impact on buyer surplus.

Table 3.9: Welfare analysis

Consumer surplus

Setting	**Average**
Communication	1.95
No communication	2.47
p-value	*0.012*

Sellers' profits

Setting	**Average**
Communication	15.27
No communication	13.63
p-value	*0.005*

3.6 Conclusion

In this paper I have examined the effects of communication and information on prices in posted offer markets with capacity constrained sellers using an experiment. In an experimental setting the level of information and the possibility to coordinate can easily be controlled, but there are also many real world examples of measures that raise the level of information or facilitate coordination: the level of information among buyers can easily be increased through information campaigns, price comparison websites or by law. Even though measures to enforce coordination are maybe less obvious, many online tools exist that help people coordinate their behavior. Facebook, for example, allows people to announce whether they want to attend an event. Furthermore, tools such as Doodle.com help people avoid miscoordination problems. Relating to the above mentioned housing market example, in many cities Facebook groups have been founded which are used by landlords or former tenants to announce empty apartments. If many people have already liked or commented on an apartment, other interested tenants may refrain from visiting the apartment as they estimate their chances of being chosen by the landlord as low when interest in the apartment is high and thus competition for the apartment fierce. These examples show that not only in experiments, but also in real world situations information and coordination can be influenced. This paper aims to find out if measures that seem beneficial at first sight will actually improve the market outcome in markets with frictions.

I find that providing buyers with the possibility to communicate fosters coordination which leads to both an increase in trade and an increase in prices. While both sellers and buyers benefit from the increase in trade, the increase in prices shifts welfare from buyers to sellers. Thus, buyers *suffer* from the possibility to coordinate. However, even with more experience, they do not stop making use of the option to communicate and mostly communicate truthfully.

I show that coordination is more likely to occur when the two sellers in a market set rather similar prices. When buyers are given the possibility to communicate, they achieve coordination by stating truthfully whether they purchase the cheaper or the more expensive product. In contrast, miscoordination follows when buyers do not communicate at all or lie about their plans.

Increasing the number of buyers with price information in a market with capacity constrained sellers provokes two opposing effects on prices: On the one hand, more buyers can observe the prices and base their purchase decisions on them. This increased overall price sensitivity induces sellers to lower their prices. On the other hand, an additional informed buyer competes with the other informed buyers for the possibility to purchase the cheaper product. As sellers are capacity constrained, buyers anticipate not being able to purchase when demand exceeds supply at a specific seller. Becoming less price sensitive they accept higher prices and thus encourage sellers to raise their prices. In my experiment, I find that the presence of communication among buyers determines which of the two effects prevails: in the treatments without communication buyers do not adjust their individual price sensitivity to the fact that more informed buyers are present. Thus, the more intuitive effect of information on prices is stronger and price are lower when more buyers are informed. In treatments with communication however, prices are higher when an additional buyer observes price information. As I do not find evidence for Lester's paradox in the treatments without communication, my experiment supports the findings of Anbarci and Feltovich (2017).

In future work I plan to investigate the effects of communication even further. One way of altering the design could be to allow subjects to send written messages instead of choosing from pre-formulated options. Furthermore, one could investigate market constellations involving only two buyers of whom one does or does not lack price information. I am curious to find out what effect communication has when sellers are not capacity constrained. Another possibility to extend the literature may be to further decrease the share of informed buyers. Instead of focusing on three buyers with at most one uninformed buyer, one might also investigate how communication affects prices in settings with at most two uninformed buyers. Additionally, it might be interesting to analyze if sellers have a positive willingness to pay for the introduction of buyer communication.

Appendix

A.1 Theoretical Background: Lester's Model

Here, I briefly describe the model this experiment is based on. For a more detailed description, see HMP (2016) or Lester (2011). In the model, capacity constrained sellers S want to sell their homogeneous products to buyers B. Suppose there is a finite number of informed buyers $N = \{1, ..., N\}$, uninformed buyers $U = \{0, ..., U\}$ and sellers $S = \{1, ..., S\}$. Buyers' maximum willingness to pay is 1 for the first unit of the good and zero thereafter. Sellers who do not incur any production costs can sell at most one unit of their good. In the first stage of the game, each seller sets his price $p_s \in [0, 1]$ and commits to this price, as renegotiations are not possible. In the second stage, each informed buyer observes the vector of prices $(p_1, ..., p_s)$ and all buyers make their purchase decisions simultaneously. Whereas uninformed buyers visit each seller with probability $1/S$, an informed buyer visits each seller with probability θ^s, where $\theta^s \in [0, 1]$ for all s and $\sum_{s \in S} \theta^s = 1$. An informed buyer wants to maximize her payoff given the prices and the strategies of all other buyers. For the informed buyer i, the probability of receiving the good if k uninformed buyers and j other informed buyers visit the same seller is $1/(j + k + 1)$. Thus, for each informed buyer, the probability of being served when choosing seller s equals:

$$\eta^c(\theta^s_{-i}) = \sum_{j=0}^{N-1} \sum_{k=0}^{U} \frac{(N-1)!}{j!(N-1-j)!} (\theta^s_{-i})^j (1 - \theta^s_{-i})^{N-1-j} \frac{U!}{k!(U-k)!} (1/S)^k (1 - 1/S)^{U-k} \frac{1}{j+k+1}. \tag{3.1}$$

In an equilibrium, informed buyers need to be indifferent between sellers such that they receive the same expected value no matter which seller they approach with positive probability. Besides, these probabilities need to add up to 1 ($\sum_s \theta^s = 1$). Together, these two conditions implicitly define $\{\theta^s(p_s, p_{-s})\}_{s \in S}$. The expected payoff from choosing seller s for a buyer is $\eta^c(\theta^s)(1 - p_s)$. Given the probability with which an informed buyer visits seller s, $(\theta^s(p_s, p_{-s}))$, and given all other prices p_{-s}, sellers maximize their expected profits in the first stage.
Therefore, the probability that a seller gets at least one buyer equals

$$\mu^c(p_s, p_{-s}) = 1 - (1 - \theta^s(p_s, p_{-s}))^N (1 - 1/S)^U. \tag{3.2}$$

Generally, both pure strategy and mixed strategy subgame-perfect equilibria exist. When the fraction of informed buyers is sufficiently small, sellers have an incentive to deviate from the pure strategy equilibrium and charge the highest possible price, in our case 1. In the experimentally relevant parameter constellations, however, the fraction of informed buyers is large enough such that all sellers set the same price and only symmetric pure-strategy equilibria exist for prices. To determine these equilibria, one needs to solve seller s' maximization problem

$$\max_{p_s} \mu^c(p_s, p^*_{-s})p_s \tag{3.3}$$

given that the other firms charge the equilibrium price. The purchase probability can be forecasted using the buyers' indifference condition: $\eta^c(\theta^s)(1-p_s) = \eta^c((1-\theta^s)/(S-1))(1-p^*)$. From the seller's first order condition, the equilibrium price can be derived:

$$p^* = M/[M + (1-\mu(p^*,p^*))\eta(1/S)\{N(B-1)/(S(N-1))\}] \tag{3.4}$$

with $M \equiv \mu(p^*,p^*)(\eta(1/S) - (1-1/S)^{B-1})$ and $\mu(p^*,p^*)$ reducing to $1-(1-1/S)^{N+U}$. Equation 4 implies an equilibrium price that can increase in the number of informed buyers for a given number of sellers and buyers (Lester's paradox). This paradox stems from the buyers' indifference condition which contains a trade-off between the price p_s and the probability to visit a seller θ^s. Consequently, each seller knows that setting a lower price raises the chance of getting at least on buyer. The degree of price-responsiveness depends on the number of informed buyers: With more informed buyers, the price-sensitivity is lower, because cheap sellers have become less attractive for buyers due to expected queues. For this reason, prices may rise when the number of informed buyers increases.

In the equilibrium with two sellers and three buyers, informed buyers approach sellers with probability $\theta^* = \frac{1}{2}$. Furthermore, μ^*, the equilibrium probability that each seller finds at least one buyer, simplifies to $\frac{7}{8}$. Thus Equation 1, the probability of being served, yields $\eta^* = \frac{7}{12}$. Plugging in $B = 3$, $S = 2$ and either $N = 2$ and $U = 1$ or $N = 3$ and $U = 0$ into Equation 4, one gets a predicted price of 0.6667 for the case with partial information and 0.7273 for the case with full information. These values need to be multiplied by 20 to get the predicted prices for the treatments without communication: 14.54 vs. 13.33.

A.2 Instructions: Partial Information and Communication

Part 1

Welcome to our experiment and thank you for participating. This experiment consists of two parts. This text will explain the first part, you will receive further information about part 2 as soon as you have completed the first part.

The first part consists of 35 rounds. In each round, there will be markets and you will either be a buyer or a seller on one of the markets. Which role you take, will be determined randomly in the beginning of the experiment. Your earnings, which you will be paid privately and in cash, depend on your decisions during the experiment.

In each round all participants will be split into four markets. A market consists of five participants of whom two will be sellers and three buyers. The participants who interact with you in the same market will be determined randomly in each round. Thus, you will not know the identity of the other participants in your market.
In the beginning of each round, every seller owns one unit of a hypothetical good. A buyer can purchase one unit of this good on her market. It is not possible to purchase or sell more than one unit per round. A round begins with sellers setting their prices independently. Prices need to be in the range of 0 and 20 ECUs and can be entered with up with two positions after the decimal point.
After sellers have chosen their prices, two of the three buyers in a market are informed about prices. The third buyer does not receive this piece of information. Which buyer has no price information is determined randomly in each round. All buyers decide if they want to communicate to the other buyers from which seller they intend to to purchase. After all buyers have had the possibility to announce their purchase plans, they decide independently from which seller to buy. The goods are homogeneous and differ only in their prices. If only one buyer decides in favor of a seller, she pays the price of the good and receives the good. If more than one buyer want to purchase from the same seller, one of the prospective buyers is chosen randomly and receives the good. The other buyers who have not been selected to purchase cannot buy in this round. If a seller does not find any buyers, he cannot sell his good.

Your earnings in a round are determined as follows:

- Sellers: If you were able to sell your good, your income is the price you demanded. If not, your income is 0.
- Buyers: If you were able to purchase your selected product, your income equals 20 ECUs minus the price you had to pay. If not, your income is 0.

Overview of the sequence of a round:

1. The computer randomly matches participants into markets that consist of two sellers and three buyers.
2. Sellers choose their prices.
3. After two of the three buyers have been informed about the prices, buyers have the opportunity to announce from which sellers they plan to purchase. Afterwards, all three buyers make their purchase decisions.
4. The round ends. If you are a seller, you are informed which price the other seller demanded, how many prospective buyers you could attract and how much you have

earned in this round. If you are a buyer, you are informed which prices the sellers had chosen, how many other buyers wanted to purchase from the seller as you, whether you could purchase the product and how much you have earned in this round.

Your earnings from part 1 of the experiment:
In the end of the experiment 10 of the 35 rounds are picked randomly. Only these ten rounds are payoff-relevant. Your average earnings from these ten rounds will be converted into Euros. For a seller, the exchange rate is: $ECU * \frac{2}{3} = €$. The exchange rate for buyers is: $ECU * \frac{8}{3} = €$. You receive your earnings from part 1 together with those from part two in the end of the experiment.

Part 2
In this part of the experiment, we show you five lotteries and ask you to pick the one you prefer. Each lottery leads to two potential results which occur with the same probability. Your earnings in this part of the experiment depend both on your choice and on the outcome of the lottery.
After you have selected one of the five lotteries, the computer randomly picks a number between 1 and 100. If this number is smaller than or equal to 50, you will receive the outcome in the left column of the table. If the number is larger than 50, you receive the outcome in the right column. In the end of the experiment, you receive your earnings from the lottery in addition to your earning from the first part of the experiment.

Chapter 4

Buying Goods of Unknown Value: An Experiment[*,**]

4.1 Introduction

Why do many people trust in alternative medicine although many alternative healing methods do not prove as objectively effective? How can the existence of some cosmetic products be explained, given that their effectiveness is at least doubtful? Why do people prefer to invest their money in actively managed funds instead of index funds?

The existence of markets for these, and similar, products is especially remarkable because information about their doubtful efficacy can easily be obtained. It is public knowledge that actively managed funds do not outperform index funds. You can easily come across newspaper articles stating that no scientific evidence of the efficacy of homeopathy has yet been found. However, consumers do not seem to use these pieces of information when making purchase decisions. In addition, there are other, maybe conflicting sources of information: they may have heard some casual stories about the products from their acquaintances; they might have come across marketing promises; and they might have made their own experiences with the products. Consequently, they may experience a lot of uncertainty when making a decision, feeling torn between these different sources of information. Thus, consumers often tend to behave as if they were

[*] I am grateful to seminar audiences at the University of Mannheim, the ESA meeting in Prague, the NCBEE in Aarhus, the Alhambra Experimental Workshop in Barcelona, and at the workshop "Coping with Difficult Decisions" in Düsseldorf for their valuable comments and suggestions.

[**] This chapter is based on a joint project with my supervisor Henrik Orzen. However, the version at hand is based on my own analysis.

unable to reach an informed purchase decision.

When buyers neglect potentially available information, the market success of some questionable products can be explained. However, these products may also be purchased by customers who do not have enough time to familiarize themselves with the respective markets. A customer who has suddenly become ill and needs quick help might ask her friends for recommendations on drugs. A drug may be recommended if the friend who took it was cured from a comparable illness. If she then gets better after taking medicine, the patient may unconsciously form a causal link between her recovery and the medicine. This belief in a causal relationship is not necessarily justified, because the patient might have recovered even if the medicine was in fact completely ineffective. But the patient never knows the counterfactual of what would have happened had she chosen a different drug or not taken any medicine. This lack of knowledge makes it hard for her to find out whether her recovery was related to the medicine. Hence, it might be both the lack of time and the neglect of available information that lead to the purchase of certain ineffective products.

In the paper 'The Market for Quacks' Spiegler (2006) models a market with extremely limited information. His leading example is the market for alternative medicine. Assuming that patients fully rely on anecdotes to reach a decision, he shows that markets for products which affect the patients' well-being just as much as the option not to purchase – markets for worthless products – can sustain in equilibrium. In the context of alternative medicine, an anecdote can be thought of as a story patients hear about the past success of a healer who treated other patients. Either the healer managed to cure his former patients or they remained ill. Spiegler assumes that patients gather one anecdote for each healer and for the option not to consult a healer and choose the cheapest of those with a positive anecdote. If people decide according to this heuristic rather than acting rationally, Spiegler's model predicts that demand for completely ineffective products persists in equilibrium.

In this paper, we study how people make decisions when they have only very limited information about the products they can purchase. Using a real effort task we create a laboratory environment which resembles a situation in which the decision maker suddenly needs a certain product, but knows little about the available products, or – in the context of Spiegler's model – becomes ill, needs medication and knows little about potential therapies and drugs that may cure the illness. An important feature of our experiment is that it consists of several independent rounds, i.e. a potential recovery lasts only for the respective round and the available products also differ from round to round. A 'sick' subject has two alternative strategies that might lead to her recovery: she can either buy costly 'medicine' or choose the option of not buying anything. The latter strategy is designed to resemble hoping for recovery without taking medicine. Subjects are informed about the prices of the different products. Moreover, they know one anecdote per costly product. These anecdotes are supposed to resemble stories patients hear about available drugs. The anecdotes are binary, i.e. they inform the subject whether the respective product would have been successful had she taken it earlier in another 'sick' round. After having made their purchase decisions, the subjects get to know the outcome of their decisions; the counterfactual remains unknown and a new round starts. This experimental setting allows us to observe the subjects' purchase decisions throughout the rounds and identify potentially used heuristics.

We have a twofold interest in analyzing how people make decisions in the context of very limited information: we look for potentially used heuristics in a one-shot scenario and over time when subjects become more familiar with the overall situation. Therefore, we split the experiment into two parts. In the first part subjects are confronted with the real effort task and need to make a purchase decision without knowing how the experiment will continue. Once all subjects have finished the first part, we inform them that the second part consists of at least 60 additional rounds in which they will either be 'healthy' or 'sick'.[1] The first part of the experiment comes closest to Spiegler's static model, whereas the subsequent rounds of the experiment show the effects of experience and learning. Note that we do not investigate the kind of learning a consumer may undergo in one *specific* market with the same suppliers over time. Instead, we want to find out to what extent consumers *transfer* their experiences from one market to another. We therefore confront our subjects with different market situations (i.e. different products, different numbers of products, different anecdotes and different prices) in each round. We seek to find out whether heuristics used towards the end of the experiment differ from those used at the beginning.

To better identify the heuristics and find out how prices and anecdotes influence purchase decisions, we introduce different scenarios which each subject will face in random order during the experiment. These scenarios differ in the number of products, prices, anecdotes and the correlation between the latter two. Since Spiegler (2006) assumes that people choose the cheapest option of those with a positive anecdote, we also stress-test this heuristic by considering scenarios in which products with positive anecdotes are more expensive than products with negative anecdotes. In the second part of the experiment

[1] Note that we use neutral language in the experiment and refer to the versions as difficult and easy.

we investigate the influence of experience on the decision makers' behavior. In a between-subjects design we implement three treatments which differ in the success probabilities of the costly products and of the option not to purchase which is costless. These probabilities which remain fixed throughout the experiment are unknown to the subjects. Furthermore, we avoid the term 'probability' in the instructions as we do not want to make our subjects aware of the fact that probabilities are involved. In the 25-25 treatment, the products are as effective as the option not to purchase and 'cure' subjects with a rather low probability of 25%. In the 75-75 treatment, the success probabilities of the products and the costless option are also identical, but substantially higher at 75%. Finally, in the 25-75 treatment, the outside option is less effective than the costly products (25% versus 75%). Since the first two treatments represent markets for quacks, they enable us to assess whether more experience with similar situations helps people to find out that the costly products are worthless. In the 25-75 treatment, the costly products are superior to the outside option. Here, we want to investigate if our subjects notice this fact and buy them in the long run. Overall, we seek to identify the most commonly used heuristics.

In the first part of the experiment we find that buyers prefer products with positive anecdotes as well as cheap products. Yet when all products with positive anecdotes are expensive, subjects rather purchase a product with a negative anecdote or do not purchase at all. The buying propensity is highest when products are cheap and have a positive anecdote. When subjects gain experience, anecdotes lose part of their importance for the purchase decisions, but subjects remain very price-sensitive in all treatments. Subjects are more likely to purchase in the 25-75 treatment than in the 75-75 treatment while the purchase propensity in the 25-25 treatment lies in between. Although the buying rates vary across treatments, the decision making heuristics do not differ systematically. Interestingly, we observe that the buying rates do not converge to 0 in the treatments where the costly products are as effective as the costless option. Similarly, they do not converge to 1 in the 25-75 treatment with the more effective costly products. While subjects in the 25-25 treatment and in the 75-75 treatment do not learn to avoid costly products, subjects in the 25-75 treatment purchase less frequently than expected.

The remainder of this paper is organized as follows: The following section gives an overview of both the experimental and the theoretical literature. Section 3 explains the experimental design followed by a section stating our hypotheses for the data analysis. In section 5, we present our results. The last section concludes and provides an outlook on future research.

4.2 Related Literature

4.2.1 Theoretical Literature

Our paper is inspired by Spiegler's (2006) model about the 'Market for Quacks'. Spiegler shows that markets for worthless goods can be sustained when consumers reason anecdotally and firms maximize profits in a standard rational way. If consumers or, in Spiegler's context, patients understood the market model correctly, they would realize that the costly goods are equally effective as the option not to purchase and the market would be inactive. Instead, they treat stories about the success of a product as fully informative. This deviation from rational behavior has first been modeled by Osborne and Rubinstein (1998): the S(1)-procedure. Decision makers sample each possible option once and choose the option leading to the best consequence. Hence, patients can be easily exploited by healers because they attribute the healers' in fact random successes to ability rather than chance and luck. Spiegler's model also shows that prices for worthless products increase when their effectiveness decreases. The reason for this counterintuitive result is that the patients' sample contains fewer positive anecdotes such that there is less price competition among the healers with positive anecdotes to attract patients. A larger number of healers raises demand for worthless products and thus harms patients because a larger sample is more likely to include positive anecdotes for a given effectiveness. The impact of the number of quacks on the patients' welfare is not monotone as a larger number of healers also increases competition and induces healers to lower their prices.

Directly related to Spiegler's model and the S(1)-procedure is Rabin's (2002) work which models believers in the 'Law of Small Numbers'. Both the discovery of this law and experimental evidence of it date back to the research of Tversky and Kahneman (1971). With his model, Rabin (2002) shows that people exaggerate the likelihood of a small sample of signals resembling the underlying population. When people do not know the rate at which signals are generated, they tend to over-infer from short sequences of signals: they believe that the underlying rate is more extreme than it is. Due to this bias, people are prone to believe in non-existing variation between in fact identical rates. Thus, in line with Spiegler's model and the S(1)-procedure people only observe very few anecdotes per quack but infer from these observations as if they were fully representative. They will conclude that some quacks will always be successful whereas others will never succeed. Hence, they believe that there is some variation among quacks, even when this is not true.

4.2.2 Experimental Literature

Tversky and Kahneman (1971) came aware of the 'Law of Small Numbers' and have found experimental evidence of it. Possible biases arising from this law include relying on small samples for making a decision, under-estimating the breadth of confidence intervals and trying to find explanations for deviations from expectations instead of attributing them to sample variability. Furthermore, Tversky and Kahnemann report overconfidence in early trends and in the stability of observed patterns as characteristic of believers in the law of small numbers. Besides, people believing in the law of small numbers tend to fall prey to the representativeness bias, considering small samples as representative for the whole population.

Builing on the work by Tversky and Kahneman (1971) on the representativeness bias, Grether (1980) formalized the information subjects receive in the experiment in order to reduce the danger of misinterpretation. He also introduced financial incentives. In his experiment, an urn – urn X – is designed to determine the subjects' prior beliefs about the likelihood of two further urns – urn A and urn B – from which the sample is drawn. Grether finds that subjects put too little weight on their prior information from urn X. When a sample is representative of either urn A or B, they overestimate the probability that the sample was drawn from that urn, although urn X attributed a rather small prior probability to that urn. However, they do not ignore the prior information completely. More experience and financial incentives can diminish the representativeness bias and induce subjects to update their information in a way that approaches Bayesian updating. More recent evidence of the representativeness bias is found in experiments by Holt and Smith (2009) who manage to avoid risk aversion issues using the Becker-DeGroot-Marshak procedure (Becker et al., 1964).

Another model of learning that may be relevant for the dynamic part of our experiment is reinforcement learning which is addressed by Roth and Erev (1995) and Erev and Roth (1998). In reinforcement learning models, the likelihood of choosing an action increases if this action has been successful in the past. The researchers find that reinforcement learning predicts decision making behavior better than rational equilibrium behavior, but augmenting the model with forgetting and experimentation improves the fit to empirical data. Charness and Levin (2005) investigate experimentally whether Bayesian updating or reinforcement learning better describe decision making under risk and uncertainty. They consider reinforcement the simpler, more natural approach. If predictions of reinforcement learning and Bayesian updating clash, decisions are inconsistent with Bayesian updating and can only be explained by reinforcement learning.

Finally, the literature on multi-armed bandits is also related to our work (for an overview, see Bergemann and Valimaki, 2006). A multi-armed bandit problem is the situation a gambler faces when confronted with a row of slot machines. The decision she has to make is which 'arm' of the multi-armed bandit to play and when to switch to a different 'arm'. Each 'arm' yields a random draw from an unknown distribution specific to that particular 'arm'. The gambler wants to maximize the sum of rewards solving the trade-off between exploiting the information he has already gathered about the arms and exploring new arms, which is costly. The situation in our experiment is similar to a bandit problem: subjects have to decide if they want to purchase and, if so, which product to choose if several products are available. The outcome of their purchase decision is a random draw from an unknown distribution. Nevertheless, there are certain differences between our work and the literature on multi-armed bandits, e.g. in our case, all costly products are equally effective, but this fact is unknown to our subjects and has to be discovered. Furthermore, subjects face a new set of costly products in each round. Additionally, we provide information in the form of anecdotes. Thus, subjects also need to learn how to deal with this source of information. Finally, our time horizon is finite which also differentiates our setting from a typical multi-armed bandit problem.

In addition to the theoretical literature on multi-armed bandit problems, there is a large amount of literature on experiments related to those problems that connects vaguely to our experimental setting. As mentioned above, a trade-off between exploitation and exploration exists: decision makers can exploit the information they have already gathered or invest in additional exploration before they make a new decision. Anderson (2012) investigates people's distaste for uncertainty about the average quality of each alternative. People tend to experiment too little because they undervalue information gained from experimentation and prefer to stick to the arm of a bandit they already know considerably well exploiting information already gathered. In our case, this would mean that subjects use a strategy of either always or never buying for many rounds. However, we find the opposite in our data: subjects experiment a lot, even in later rounds.

Banks et al. (1997) design an experiment with two two-armed bandits, one of which has one certain and one uncertain arm, whereas the arms of the other bandit are both uncertain. They want to test if behavior is more myopic when subjects face the bandit with two uncertain arms than when choosing between one certain and one uncertain arm. In the latter case, subjects should always choose the uncertain arm when the expected payoff of both arms is identical in order to gather more information about the arm. The observed behavior comes close to their expectation. In their setting, time is infinite and the identity of the bandits does not change over time, whereas in our design, the opposite holds true. In addition, we did not include a certain arm in our experiment.

Norton and Isaac (2012) also treat ambiguity aversion in the context of bandit problems. Their setting can be interpreted as an expert-client-relation: in their experiment, experts with a conflict of interest try to convince clients to switch to an unknown arm. The other arm yields a certain payoff. While the client does not know her payoff of the unknown arm before selecting it, the expert is informed about it and may lie to the client about its value in order to convince her to switch to this arm. Thus, ambiguity stems from the fact that the expert's recommendation is not necessarily trustworthy. Clients react to this ambiguity by mostly avoiding the unknown arm although switching might have been beneficial for them. In our experiment, the supply side of the market is computerized and only the demand side of the market is taken over by subjects. Thus, we exclude conflicts of interest. In our setting the fact that both the effectiveness of the products and the effectiveness of the costless option are unknown creates ambiguity. Furthermore, the identity of the products changes such that it is unclear for subjects if they can benefit from previously gathered observations on the performance of the products when deciding between purchasing and not purchasing in the current round.

4.3 Experimental Design

The experiment was programmed in VisualBasic and it was run at the University of Mannheim in the mLab in Spring 2014. Subjects were recruited through ORSEE (Greiner, 2004). The majority of our 111 participants were undergraduate students from all fields. The experiment consists of three between-subject treatments in which we alter the effectiveness of the costly products and of the option not to purchase. At the beginning of the experiment subjects play five training rounds to familiarize themselves with the real effort task. Afterwards all subjects participate in the first payoff relevant part of the experiment which is designed to resemble a one-shot decision making situation. As soon as all subjects have completed the first part of the experiment, they receive new instructions and enter the second part of the experiment which consists of at least 60 rounds.

4.3.1 Real Effort Task

In each round subjects need to solve a real effort task. If subjects solve the real effort task successfully, they earn 100 points (€5). We implement two versions of the real effort task, one easy version which subjects will most likely be able to solve successfully and one difficult version which is almost unsolvable. In order to avoid biases caused by the context, we do not use any specific framing for the task in the laboratory and in the instructions. Instead we refer to the versions as 'easy' and 'difficult'. If subjects are confronted with the difficult version of the real effort task, they can purchase potentially effective remedies that possibly transform the difficult version of the real effort task into an easy one. However, even without purchasing help, the difficult version of the real effort task potentially turns into the easy version, but perhaps it stays difficult. Although the difficult version is transformed into the easy version with a treatment-specific probability, we do not mention the word 'probability' or other terms which subjects might relate to statistics. The vague formulations add to the realism of our setting as decision makers do not necessarily know that probabilities might be involved in naturally occurring situations.[2]

The real effort task is designed to abstractly resemble Spiegler's motivating example of unexpectedly becoming sick. Lacking time, knowledge or motivation to familiarize themselves with all potentially effective drugs and treatments, patients still need to select one or decide to hope for recovery without taking medication. Both strategies may be effective. In the experiment the difficult version of the real effort task resembles being sick and it is designed to create a positive willingness to pay for potentially effective products.

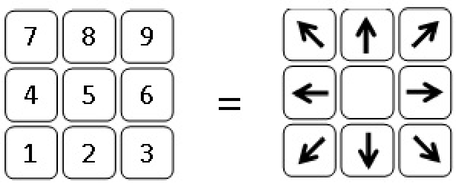
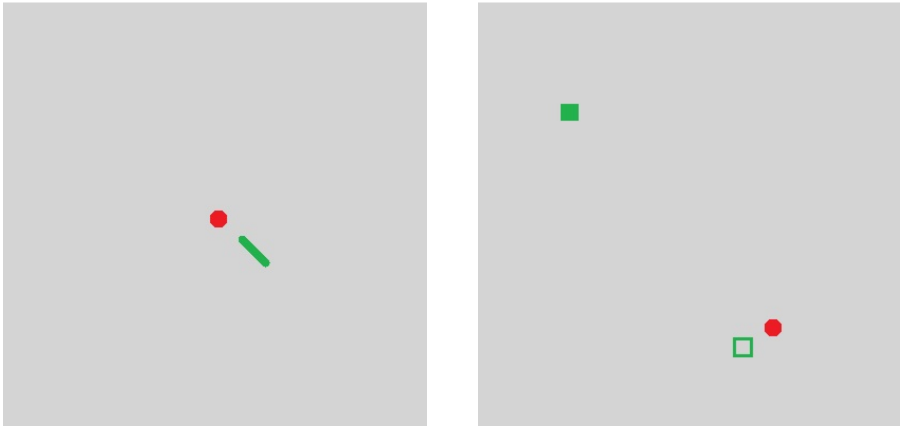


Figure 4.1: Steering the cursor

The real effort task consists of a game in which subjects have to find and hit a number of targets hidden on the screen. In order to do this they are instructed to steer a red ball through a two-dimensional field using the NumPad of the keyboard (see Figure 4.1). Moving the ball with the number keys is intuitive: for example, pressing '9' moves it into the upper-right corner of the screen, whereas '2' moves the ball vertically downwards. Subjects have to 'hit' the hidden targets with the ball. A small green line next to the red ball always points towards the position of the hidden target that has to be found next (see Figure 4.2a). When the ball is within a small distance of the hidden target, the outline of

[2] Note that we do not deceive our subjects, but instead keep the information which we reveal to our subjects vague.



(a) Target in the bottom right corner

(b) One target found, a second one within reach

Figure 4.2: Real effort task

the target becomes visible, i.e. a green empty square appears. Seeing the exact position of the target makes it easier for subjects to hit the target. A target that has been found, i.e. that has been hit with the ball, turns into a green filled square (see Figure 4.2b). Once a subject has found a target, the green line will point towards the next target immediately.

The two versions of the real effort task differ in the number of hidden targets: in the easy version subjects have to detect three targets within a time limit of 20 seconds, while in the difficult version seven targets have to be found. The task is only solved successfully if *all* targets have been found within the time limit. We determined the respective number of targets based on our observations from pre-tests ensuring that the easy version is doable while the difficult version can almost never be solved successfully. To allow subjects to become familiar with the real effort task, we give them the opportunity to play five training rounds of which two are easy and three difficult. This experience puts them in a position to judge the respective difficulty of both versions of the real effort task and develop a positive willingness to pay for a reduction in the number of targets from seven to three.

4.3.2 Boosters

When subjects are confronted with the difficult version of the real effort task in a payoff-relevant round, subjects are offered costly products: 'boosters'. These boosters potentially reduce the number of targets from seven to three and thus convert the difficult version of the real effort task into the easy version. However, subjects also have the possibility to refrain from purchasing by selecting the option 'Proceed without booster' which also has the potential of reducing the number of targets. If the task becomes easy after the subject has chosen a booster, a message appears on the screen notifying her that the booster has worked. In this case, only three targets have to be detected. When the booster has not worked, the subject is also informed and still has to try finding all seven targets. If the task becomes easy after the costless option to proceed without booster is chosen, the subject is notified that she has been lucky and the number of targets has been reduced to three.

The boosters are designed to resemble Spiegler's quacks. Boosters can potentially be effective, yet subjects have little information to rely on when choosing a booster or when deciding not to purchase. In Spiegler's model, patients observe one anecdote for each quack. This anecdote can be thought of as an observation on the quack's past performance. Anecdotes contain information if the quack healed his former patient or if the patient remained sick. In the model, all quacks and the option not to purchase are equally effective, i.e. they cure patients with the same probability. Anecdotes are generated through draws from an underlying probability distribution. Spiegler's central assumption is that decision makers fully rely on these anecdotes and visit the cheapest quack equipped with a positive anecdote. This decision making heuristic is referred to as the S(1)-procedure. Markets for useless products and services can exist when the demand side behaves according to the S(1)-procedure. Furthermore, Spiegler shows that prices are higher when quacks are less effective as patients observe fewer positive anecdotes. Also an increase in the number of quacks may be harmful, because the likelihood that a sample contains at least one positive anecdote increases in the number of quacks.

In order to capture this idea in our experimental design each booster is equipped with a binary anecdote. For a positive anecdote, a box next to the name of the booster is checked. In case of a negative anecdote, the box is unchecked.[3] We inform our subjects that a checked box means that the booster would have been able to reduce the number of targets whereas boosters with a negative anecdote would not have converted the difficult version into the easy version in the previous round. We point out explicitly that anecdotes do not inform subjects if the booster will be successful again in the current round.

The probabilities with which boosters and the option not to purchase reduce the number of targets vary across treatments. They are identical to the probabilities with which a positive anecdote is generated in each treatment. In the 25-25 treatment boosters and the option not to purchase reduce the number of targets in 25% of the situations. Accordingly, 25% of the anecdotes are positive and the remaining 75% negative. In

[3] For an illustration, see the instructions in the appendix.

the 75-75 treatment the number of targets is decreased with a probability of 75% when a booster is purchased or when the subject does not purchase. Only in the 25-75 treatment does the success probability (75%) exceed the success probability of the option not to purchase (25%). Remember that both the exact probabilities and the fact the probabilities are involved are unknown to our subjects.

In Spiegler's model quacks can differ in their anecdotes and in their prices. Since the anecdotes result from their effectiveness which is exogenously determined by chance, prices are the sellers' only choice variable. As we are interested in observing the purchase behavior, subjects take the role of buyers while our boosters are computerized. In addition to the anecdotes, subjects observe the prices of the boosters. In the experiment the prices of the boosters can vary between 0 and 50 points. For a better overview of the available boosters, subjects have the possibility to sort the boosters according to their names, their anecdotes and their prices. In each round, subjects are confronted with new boosters, i.e. they never interact with the same booster twice. We inform subjects about this fact and tell them to consider each round separately. By changing the boosters' identity from round to round we ensure that potential learning applies to the situation in general, not to the experience with one specific setting.

In order to better observe the subjects' decision making heuristics, we confront subjects with different constellations of boosters: we vary the number of available boosters, their anecdotes and the boosters' prices systematically (see Table 4.1). 'Low' prices are chosen from the interval $\{1, 25\}$, whereas 'high' prices are picked from the interval $\{26, 50\}$. We include all possible price and anecdote combinations with one booster to observe the influence of prices and anecdotes on the subjects' propensity to purchase (Scenarios 1 – 4). To analyze how prices affect the subjects' choice between boosters, we include two scenarios with boosters of the same anecdote, but different prices (Scenario 5 and 6). Similarly, in order to investigate our subjects' understanding of anecdotes, we implement boosters with identical prices, but different anecdotes in Scenario 8. Scenario 7, 9 and 10 were designed to test if subjects select the cheapest booster equipped with a positive anecdote, i.e. if they employ the S(1)-procedure. As Spiegler reaches the result that markets for useless products exist by assuming that buyers behave according to the S(1)-procedure, we want to find out if this heuristic is actually applied. Scenario 7 and 9 are designed to stress-test the S(1)-procedure: in these scenarios all boosters with positive anecdotes are expensive. Thus, the purchase behavior in these scenarios will reveal if subjects still purchase according to the S(1)-procedure when cheaper options with negative anecdotes are available. Hence, we can find out if subjects value positive anecdotes more than low prices. Each subject faces each scenario several times, yet the likelihood to

Table 4.1: Overview of scenarios used

Scenario	Number of boosters	Anecdotes and prices
1	1	Positive anecdote: Low price
2	1	Positive anecdote: High price
3	1	Negative anecdote: Low price
4	1	Negative anecdote: High price
5	2	2 positive anecdotes: Low price vs. high price
6	2	2 negative anecdotes: Low price vs. high price
7	2	Positive high-price vs. negative low-price anecdote
8	2	Pos. vs. neg. anecdote: Identical (low) prices
9	20	Positive high-price, negative low-price anecdotes
10	20	Pos. and neg. anecdotes: Uncorrelated prices

Note: Low price = chosen from 1,. . . ,25; high price = chosen from 26,. . . ,50

be confronted with a specific constellation depends on the treatment. In the 25-25, for example, subjects are more likely to be offered one booster with a negative anecdote (Scenario 3 and 4) than in the 75-75 treatment, where Scenario 1 and 2 are more likely to occur.

4.3.3 Timing

At the beginning of the experiment subjects receive the instructions for the five practice rounds and for the first part of the experiment, but not for the second part of the experiment. However, they are informed that a second part will follow. After completing the last practice round, all subjects are confronted with the difficult version of the real effort task facing one of the ten scenarios from Table 4.1. Each subject receives an endowment of 100 points which she can spend on one of the available boosters. If she manages to solve the real effort task successfully – independent of the version – she earns 100 additional points. Subjects have to decide if they want to spend part of their initial endowment on a booster or if they want to keep their endowment and hope that the difficult version of the real effort task will be converted into the easy version. The payoff from the first part of the experiment consists of the initial endowment of 100 points minus the points spent on a booster if a booster is purchased plus 100 additional points if the task is completed successfully.

Once all subjects have completed the first part, the experiment is paused and subjects receive the instructions for the second part. The second part of the experiment consists of at least 60 rounds. It ends when the slowest subject in a session has completed 60 rounds. In each round subjects have to play the real effort task. Which version of the task subjects are confronted with varies on a random basis. However, we ensured that ten of the 60 real effort tasks are easy. Whenever subjects face the difficult version, they are offered boosters. Hence, the difficult rounds in the second part resemble the first part of the experiment. The possible scenarios are the same as in the first part of the experiment (see Table 4.1), yet in the second part each subject is confronted with the each scenario several times. Subjects encounter the various scenarios in frequencies such that the sample information from the anecdotes is overall representative of the relevant treatment (25% or 75%). As in part 1, each subject receives an endowment of 100 points at the beginning of each round. If the round is difficult, she can spend part of her endowment on a booster. She earns 100 additional points if she solves the real effort task successfully.

4.3.4 Payoff

Two rounds of the experiment are payoff-relevant: the round played in the first part and one randomly determined round from the second part. Thus, subjects who solve both relevant real effort tasks successfully without spending any points on boosters in these rounds earn 400 points (20 €). On average, our subjects earned 9.32 €.

4.4 Hypotheses

What kind of decision-making heuristics do we expect to see in our data? Can we find evidence of the S(1)-procedure? Or are other heuristics more prominent in subjects' behavior? Additionally, how does the subjects' behavior change when they gain experience? In order to structure the data analysis, we postulate the following hypotheses, which we want to analyze separately for the first and the second part.

H1a When purchasing, people buy the cheapest product out of those with a positive anecdote (S(1)).

When seeing a positive anecdote, people who behave according to S(1) will believe that the booster belonging to the anecdote is always successful. People thus over-infer from sample information. We look for evidence of this behavior in the data. To stress-test this principle, we created two scenarios where all boosters with positive anecdotes were expensive. We compare the behavior in both kinds of scenarios to find out if the S(1)-procedure is stronger than the attraction towards cheap boosters.

H1b. When purchasing, people buy only products with positive anecdotes.

This hypothesis constitutes a softer version of Hypothesis 1a, as it makes no statement about the prices at which people buy; only the selected anecdotes are considered. Especially in the first round, we expect people to rely on the anecdotes because they are the only source of information they have. Over the rounds, they may learn that anecdotes do not help much to predict effectiveness.

H2 When purchasing, people buy the cheapest good regardless of its anecdote.

Hypothesis 2 suggests that subjects focus on prices instead of anecdotes as an alternative purchase criterion. We suspect that subjects are very price sensitive because they are provided with almost no information about the quality of the boosters. However, there is evidence that people regard the price of a product as an indicator of its quality (see e.g. Jacoby et al. (1971) for evidence related to beer prices and Acebrón and Dopico (2000) for beef prices). In contrast to these papers, marketing is not an aspect of our research, as we do not make any promises to subjects concerning the relationship between price and efficacy. Still, it does not seem inconceivable that previous situations our subjects have been in have taught them that price and quality are positively correlated. But will people also consider price as an indicator of quality in our setting and prefer more expensive products? Will their behavior change over time?

H3 People buy equally often in the 25-25 treatment and the 75-75 treatment.

The first two treatments are identical in the sense that boosters are as effective as the outside option; it is only the success probability that differs. Thus, we expect to observe similar behavior in the first two treatments during the last rounds of the experiment. However, it could well be that the frustration subjects experience in the 25-25 treatment induces them to behave differently than in the 75-75 treatment. Possibly, subjects experiment more in the first rounds of the 25-25 treatment because the frequent disappointments with both the boosters and the costless option tempt them to switch their buying strategy frequently. In the 75-75 treatment, subjects experience success more frequently and are thus less likely to change their strategy once they have found a satisfying strategy. However, with more experience subjects should behave similarly in both treatments.

H4 In the 25-25 treatment and the 75-75 treatment, the share of people purchasing boosters converges to 0.

These treatments are identical in the sense that boosters are as effective as the costless option, thus people should stop buying costly boosters in the long run.

H5 In the 25-75 treatment, the share of people buying boosters converges to 1.

In this treatment, boosters are more effective than the costless option. We expect subjects to become aware of this fact and to develop a willingness to pay of up to half of their initial endowment for a booster: in this treatment boosters simplify the real effort task with an underlying probability of 75%. Even though subjects do not know this probability nor the fact that probabilities are involved, over the rounds they may develop the impression that boosters are superior to the option of not purchasing. Observing that on average 75% of the purchased boosters have been successful, they may expect to earn about 75 points when they purchase. The costless option is less often successful, such that subjects should attribute a lower expected value, 25 points in expectation, to this option. The difference in the expected values determines their willingness to pay for a booster. On average, subjects earn 50 points more when purchasing than when choosing not to purchase. As no price in the experiment exceeds 50 points, experienced subjects should always purchase a booster.

4.5 Results

4.5.1 Task Performance

We begin our analysis by taking a brief look at the participants' performance in the real-effort task. For the purpose of our experiment it is important that our subjects attach a value to becoming 'healed'. Therefore, we sought to calibrate the difficulty of the tasks in such a way that players could expect to succeed in the easy 3-target task but fail in the harder 7-target task. Figure 4.3 suggests that our calibration worked well, by and large: Success rates for the hard task are indeed near zero whereas success rates for the easy task converge to 100% over time. There appears to be some learning though: In the last practice round and in part 1 there were still some, about 20%, who failed to complete the easy task successfully.[4] However, most of those who

[4] The last practice round in fact always featured a hard task. In Figure 4.3 we report the proportion of subjects who had managed to hit at least three of the seven targets within the 20 second time limit.

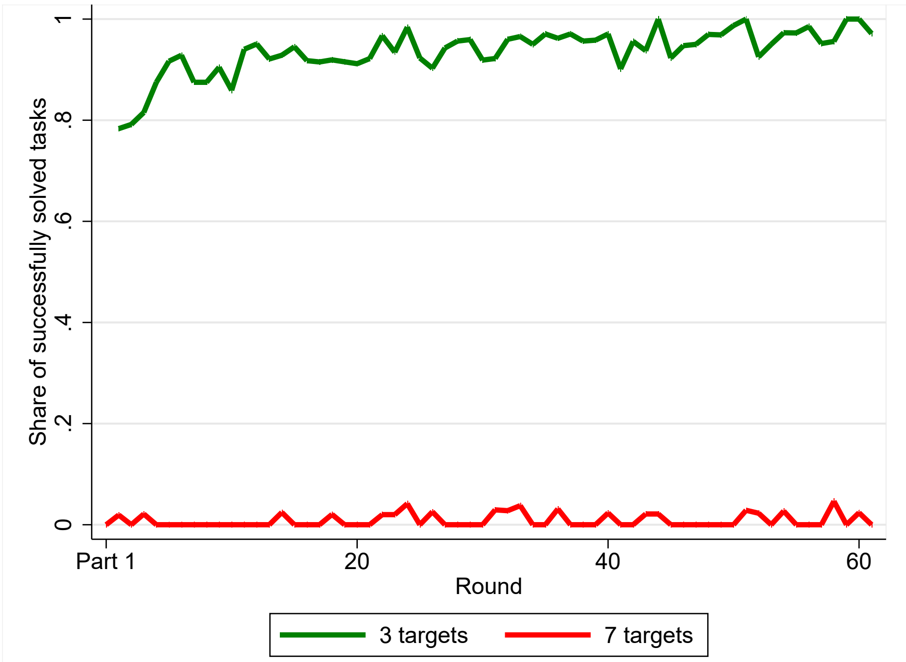


Figure 4.3: Task success rates

did not succeed in these rounds managed to hit at least two targets within the time limit.[5]

Based on these results one would not expect that the maximum booster price of 50 exceeds a subject's induced willingness to pay for a reduction in the number of targets from seven to three. Nevertheless, we do find that poor performance in the last practice round (hitting fewer than three targets) lowers the probability that a subject will buy in part 1 (p-value = 0.003; McNemar's test). However, in later rounds we do not detect any relationship between task skill and buying propensity: the part 2 buying rates of those who performed well and those who performed poorly in the last practice round or in part 1 do not differ significantly. There is also little correlation between task performance in rounds 1-30 and buying rates in round 31-60 (Spearman rank correlation coefficient = 0.113; p-value = 0.241).[6]

4.5.2 Part 1

The one-shot nature of our part 1 comes closest to the static model in Spiegler's 'Market for Quacks'. Thus, this is perhaps the most suitable setting to study the empirical relevance of the S(1)-procedure. In this subsection we will therefore focus on part 1. In

[5] 17 of the 26 subjects who hit fewer than three targets in the last practice round hit two targets and 20 of the 25 subjects who hit fewer than three targets in part 1 hit two targets.

[6] A subject's "task performance" here is defined as the subject's relative frequency of successfully completing the easy task (3 targets).

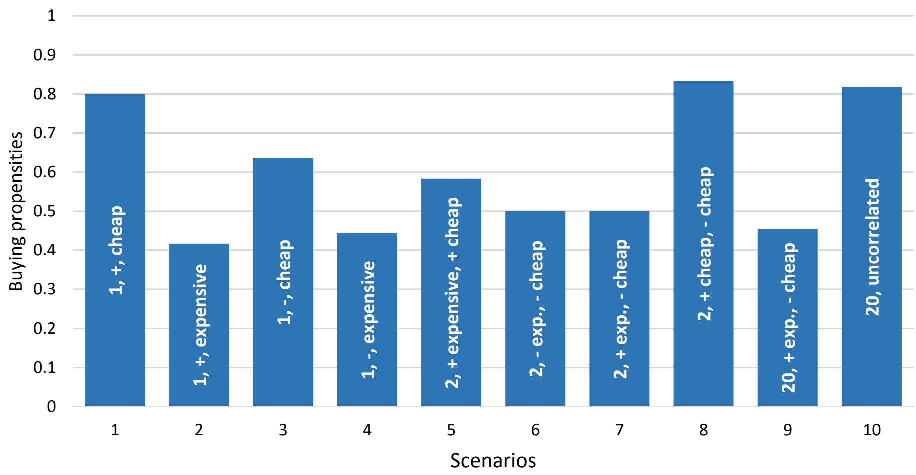


Figure 4.4: Buying propensities by scenario

Section 4.5.3 we will examine part 2 and in particular the behavior of experienced players.

First of all, we want to know if subjects buy boosters the first time they are confronted with them. Furthermore, we would like to find out in which scenarios they are most likely to purchase. Figure 4.4 depicts the purchase propensities for each scenario. Subjects are most likely to purchase in Scenario 1 when there is only one cheap booster with a positive anecdote, in Scenario 8 with two cheap boosters of which one has a positive and the other one a negative anecdote and in Scenario 10 where they can choose among 20 boosters of which some cheap boosters have positive anecdotes. The purchase propensity is lowest when subjects are confronted with only one expensive booster (Scenario 2 and Scenario 4). In these situations the anecdote does not seem to matter. Similarly, subjects are reluctant to purchase when there are many boosters, but all of those with a positive anecdote are expensive (Scenario 9). From Figure 4.4 it seems as though high prices are not interpreted as signals for quality, but instead discourage subjects from purchasing.

Focusing on the buying rates in the scenarios with only one booster (Scenarios 1 – 4), the buying propensity is highest when the anecdote is positive and the price is low. Switching to a negative anecdote reduces the buying propensity only a little (from 0.80 to 0.64) and this change is not significant (p-value = 0.635, Fisher's exact test). Increasing the price has a larger impact and the buying propensity falls to 42% (p-value = 0.082). When the price is high, the anecdote appears not to influence buying propensities at

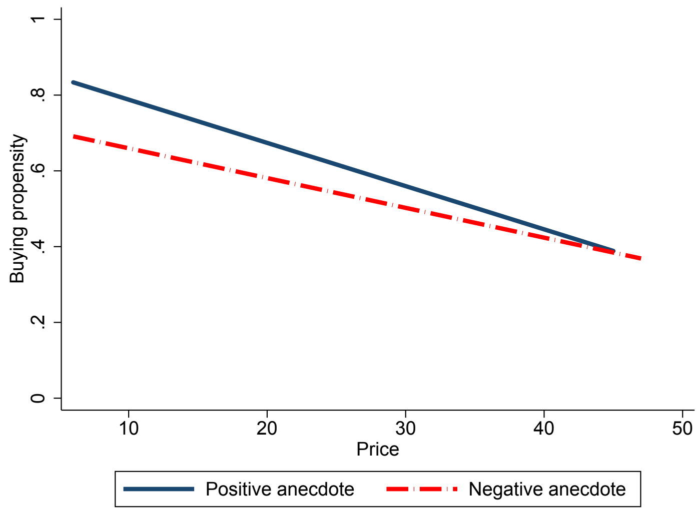


Figure 4.5: Regression: Purchase propensity

all. A simple linear probability model[7] seeking to explain the purchase decision with the price of the offered boosters and an anecdote dummy which equals one when the booster's anecdote is positive confirms these results: the price coefficient is marginally significant (p-value = 0.086) while the anecdote coefficient is not (p-value = 0.633). Figure 4.5 shows the estimated regression lines.

Result 1: In scenarios with only a single booster on offer, subjects are most likely to buy when the booster has a low price and a positive anecdote, consistent with the S(1)-procedure. However, buying rates depend more on price than on anecdote.

So how do subjects interpret our implementation of positive 'anecdotes' in the presence of several boosters? Are they seen as a signal of quality, are they ignored, or are they even taken to mean something bad? Our data suggest that ceteris paribus most subjects prefer boosters with a positive anecdote to negative-anecdote boosters when given the choice. Consider Scenario 10, where 20 boosters are available and anecdotes and prices are uncorrelated, and Scenario 8, where two identically priced boosters feature different anecdotes. For Scenario 10 we find that 78% of boosters purchased have positive anecdotes. Similarly, in Scenario 8 80% of those who decide to purchase a booster buy the one with the positive anecdote. Under the null hypothesis that boosters with positive and negative anecdotes are equally likely to be chosen a two-sided binomial test applied to this data delivers a p-value of 0.019. Thus, the data provide support for our hypothesis

[7] Standard errors are clustered on subject-level.

that subjects prefer boosters with positive anecdotes (see Figure 4.6).

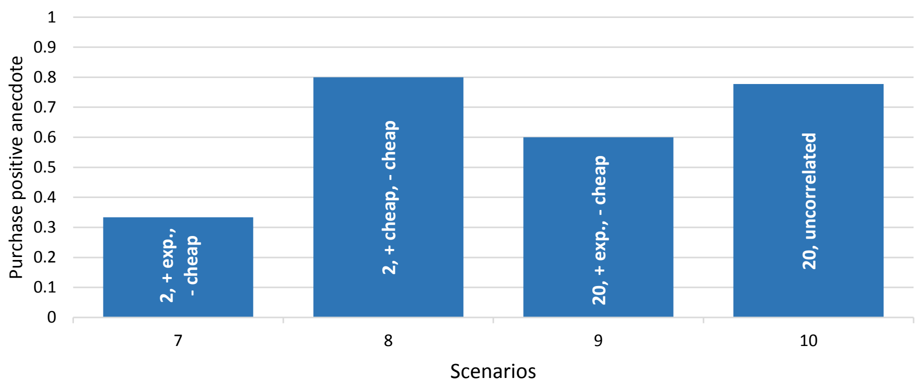


Figure 4.6: Propensity to purchase a product with a positive anecdote

Are prices interpreted as a signal of quality? As conjectured before, the data rejects this hypothesis very clearly. In fact, none of the subjects facing Scenario 10 bought a high-price booster. Likewise, all subjects facing Scenarios 5 and 6 (two boosters with identical anecdotes) opted for the cheaper booster when they made a purchase supporting our second hypothesis.

Putting these findings together, there is some indication of S(1)-like behavior. In fact, in Scenario 10 – in which subjects have the choice between any combination of low- versus high-price and positive- versus negative-anecdote boosters – 67% of all purchases are precisely consistent with the S(1)-procedure and 78% are roughly consistent with the S(1)-procedure.[8] In summary, we obtain

Result 2: Ceteris paribus, subjects opt for low-price boosters and tend to choose boosters that are associated with positive anecdotes.

However, does the S(1)-procedure still predict as well when positive-anecdote boosters are more expensive than negative-anecdote boosters? The answer is no. In Scenario 9 subjects can choose among 20 boosters where those with a positive anecdote have a

[8] A purchase decision is roughly consistent with the S(1)-procedure if the cheapest or the second cheapest booster was bought.

high price and those with a negative anecdote have a low price. In our data for this scenario we have *no* observation in which a subject chooses precisely according to the S(1)-procedure although 20% of purchases meet the criteria of rough consistency. It is noteworthy that a majority of subjects does not buy at all now (55%, compared to just 18% in Scenario 10). In Scenario 7 (one low-price negative-anecdote booster versus one high-price positive-anecdote booster) only 17% of the subjects choose to buy the expensive alternative. Again, many subjects do not buy at all (50%, compared to only 16% in Scenario 8) and some (33%) buy cheap negative-anecdote boosters. Thus, in Scenario 7 one third of all purchases are consistent with the S(1)-procedure.

Result 3: When positive-anecdote boosters are expensive subjects tend to not buy them.

4.5.3 Part 2

The role of part 2 in our design is to give us an opportunity to study the choices of experienced consumers in a 'Market for Quacks' kind of setting. This experience, however, is not tailored towards one particular market. We begin this part of the analysis by investigating how aggregate buying behavior evolves over time in the three treatments. Subsequently, in Section 4.5.3 we will fit a reinforcement learning model to our data and in Section 4.5.3, we will once again consider the different scenarios separately and reexamine our findings from part 1 after subjects have had experienced vastly different success rates across treatments.

Buying behavior over time

Figure 4.7 is a time series graph of the buying propensities (the number of instances where a booster was bought divided by the number of instances where boosters were offered) in our three treatments over the course of the experiment. This provides us with a first glimpse at learning in this environment: to what extent do subjects learn to avoid or seek boosters? As Figure 4.7 suggests, there appears to be relatively little development in buying propensities over time. Specifically, within the second half data (rounds 31-60) we detect no systematic time trends in any of the treatments. However, relative to part 1 buying rates in the second half of the experiment are significantly lower in the 25-25 treatment and in the 75-75 treatment (p-values of 0.07 and 0.05, respectively)[9]. In the 25-75 treatment the increase relative to part 1 is not statistically significant (p=0.16). Nevertheless, buying rates remain high in the 25-25 and the 75-75 treatments even after a large number of repetitions and clearly do not converge to 0 as we expected. Subjects

[9] The reported p-values are based on two-sided one-sample Fisher-Pitman permutation tests.

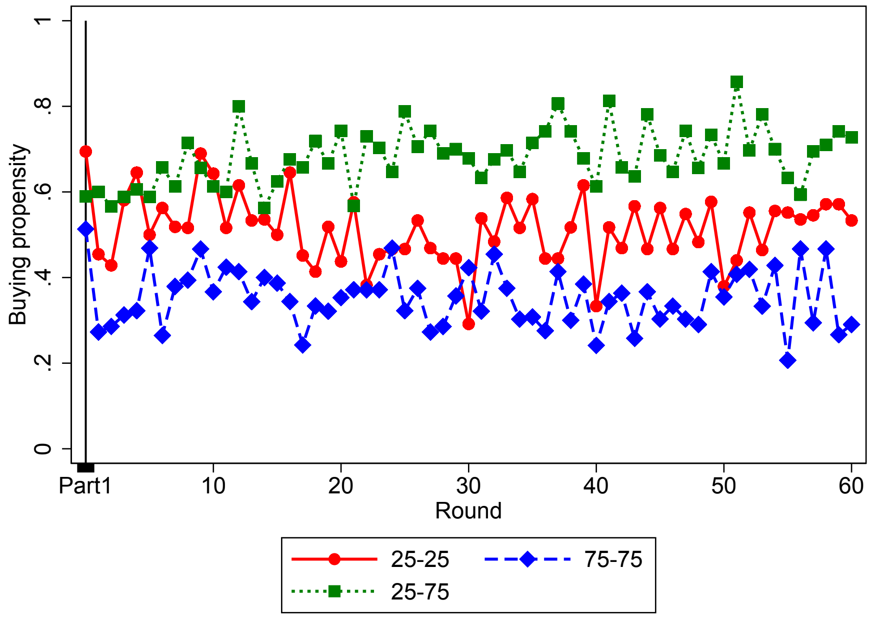


Figure 4.7: Buying propensities over time

do not seem deterred by the poor reliability of the boosters in the 25-25 treatment or very attracted by the good performance of the outside option in the 75-75 treatment.

In the 25-75 treatment the buying propensity does not grow much above 70% even towards the end of the experiment which contradicts our hypothesis of a buying rate converging to 1. Thus, despite the substantial performance difference between boosters and outside option subjects do not appear to develop a habit of acquiring boosters routinely.

Notwithstanding the overall mild dynamic adjustments, buying rates in the second half differ significantly across treatments. A two-sided two-sample Fisher-Pitman permutation test delivers a p-value of 0.035 for the comparison between the 75-75 and the 25-25 treatment and a p-value of 0.019 for the comparison between the 25-25 and the 25-75 treatment.

In order to identify further influencing factors on the purchase decision, we regress the propensity to purchase on the price of the cheapest available booster and the price of the booster which should be chosen if subjects behaved according to the S(1)-procedure (see Table 4.2). These two variables capture the subjects' price sensitivity. Furthermore, we include a gender dummy, a dummy which equals one in the 25-75 treatment, i.e. in the treatment with superior boosters, and a dummy for the 75-75 treatment. We control for experience by including the round number. We use two different models while clustering the standard errors on subject level: in addition to the linear probability model which facilitates the interpretation of the coefficients, we also provide the outcomes of a random effects panel logit regression in Table 4.2. In line with the visual impression from Figure 4.7 we find that the round number does not have a significant impact on the propensity to purchase. Furthermore, a lower S(1)-price and a lower minimum price increase the subject's likelihood to purchase significantly. In the 25-75 treatment subjects are significantly more likely to purchase, while in the 75-75 treatment subjects are less likely to purchase than in the 25-25 treatment. Female subjects purchase less often. The reason for this observation might be that female subjects are more risk-averse and therefore do not want to spend their endowment on boosters of a doubtful effectiveness.

Table 4.2: Regression: Buying propensity in part 2

	(1) Linear Probability	(2) Logit
Min. price	-0.007***	-0.078***
	(0.001)	(0.006)
S(1)-price	-0.002***	-0.026***
	(0.000)	(0.005)
25-75 treatment	0.172**	1.744**
	(0.077)	(0.823)
75-75 treatment	-0.169**	-1.826**
	(0.078)	(0.825)
Female	-0.126**	-1.424**
	(0.063)	(0.669)
Round	0.000	0.003
	(0.000)	(0.003)
Constant	0.747***	2.634***
	(0.063)	(0.681)
(Pseudo) R^2	0.1688	0.4011
N	4138	4138

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered on subject-level.

Buying propensities are not at all homogeneous across subjects. Indeed, some of them buy in almost every round, others hardly at all and yet others buy occasionally or often. Figure 4.8 shows histograms of buying propensities in the three treatments. In the 75-75 and 25-75 treatments, we obtain uni-modal distributions with peaks in the expected locations (rare buying in 75-75 and frequent buying in 25-75). In the 25-25 treatment, there is no clear pattern and a majority of participants appears to be rather undecided. We suspect that this may be a result of the fact that both options are characterized by low healing rates: subjects may be inclined to continue to experiment and search for the better alternative out of sheer frustration.

Result 4: Learning appears to be difficult in our setting. In spite of quite extensive experience many subjects continue to buy even when boosters are ineffective, while others continue to use the outside option even when boosters are effective.

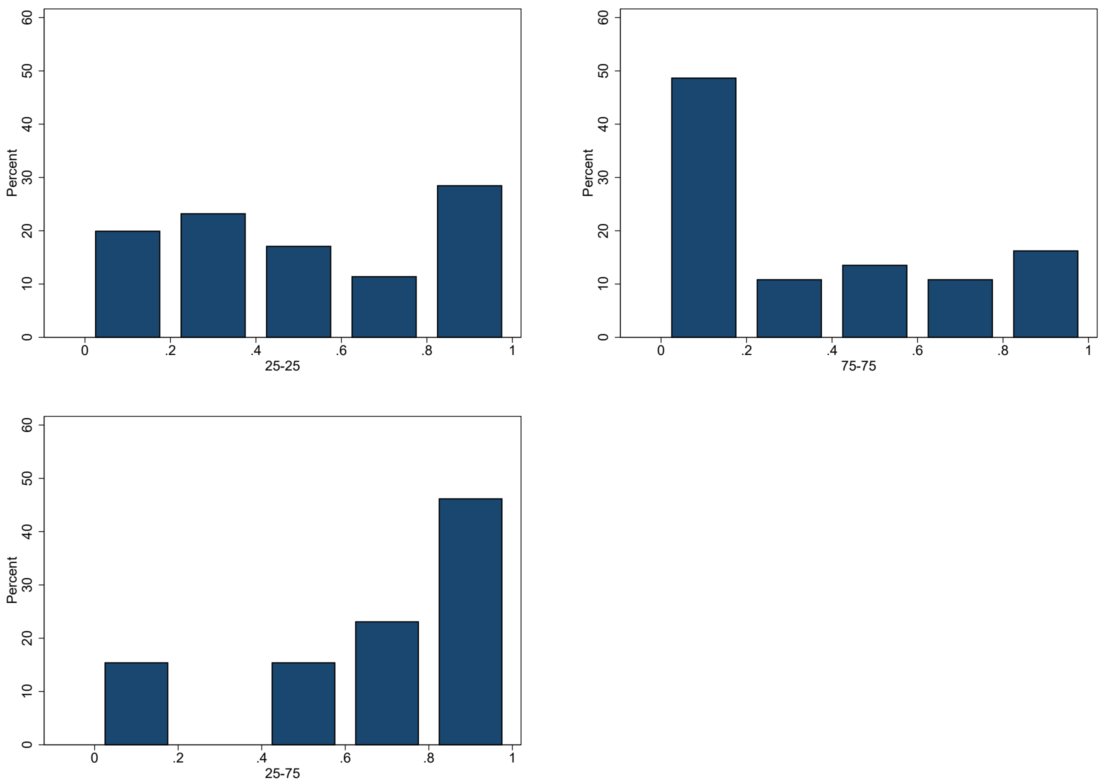


Figure 4.8: Distribution of individual buying propensities

Reinforcement learning

To further study the observed behavior we fit a reinforcement learning model to our data. We consider reinforcement learning a very natural approach in our setting: Bayesian learning is often too demanding for subjects as it requires a sound understanding of the situation. Furthermore, Roth and Erev (1995) find that Bayesian updating is inferior to models of reinforcement learning in predicting experimental behavior. Building on their results, we ask if reinforcement learning can explain the observed behavior in our experiment. Do subjects use profitable strategies with increasing probability? To apply the idea of reinforcement learning to our experiment, we need to define the possible strategies subjects could choose. As subjects faced new boosters and new scenarios in each difficult round, we cannot track behavior towards specific boosters. Instead we focus on the decision whether or not to purchase. We also abstract from potentially recurring patterns in the buying decisions such as always choosing the cheapest booster or never purchasing a booster with a negative anecdote.

In reinforcement learning models, the probability with which a strategy is played is derived from the propensities of all possible strategies. The propensity R_{ikt} to choose strategy $k \in \{\text{buy}, \text{not buy}\}$ in the current round depends on the last round's propensity R_{ikt-1} and on the last round's payoff π_{ikt-1}. The payoff π_{ikt-1} denotes the payoff earned when the decision maker i chose strategy k in the previous round. The parameter q captures the influence of last round's propensity to choose strategy k on the current propensity. Thus, $q_i \in [0, 1]$ measures how much subjects discount past information.

$$R_{ikt} = q_i R_{ikt-1} + x\pi_{ikt-1} \tag{4.1}$$

In Equation 4.1, $x \in \{0, 1\}$ is an identifier which equals 1 if strategy k was chosen in the previous round, otherwise it is 0. The propensity to select a strategy is computed for both possible strategies k, i.e. for purchasing a costly booster and for not purchasing. If a strategy was not employed in the previous round, its propensity is discounted by q_i and no additional payoff π is added. For the strategy to purchase a costly booster the price of the chosen booster needs to be subtracted no matter if it simplified the task and 100 points could be earned or not. The payoff of selecting the costless option equals 100 if the costless option worked or 0 as no costs need to be subtracted.[10]

$$\pi_{\text{buy}} = \begin{cases} 100 - p & \text{if booster worked} \\ -p & \text{if booster did not work} \end{cases} \tag{4.2}$$

$$\pi_{\text{not buy}} = \begin{cases} 100 & \text{if costless option worked} \\ 0 & \text{if costless option did not work} \end{cases} \tag{4.3}$$

Based on the propensity to choose the costless option and on the propensity to purchase the probabilities with which the strategies are implemented can be computed. Here, we follow the approach of Chen and Tang (1998) and use an exponentialized functional form.

$$p_{ikt+1} = \frac{e^{\lambda R_{ikt}}}{\sum_k e^{\lambda R_{ikt}}} \forall k \tag{4.4}$$

The parameter λ captures the influence of the propensities on the probabilities with which each strategy is chosen. When $\lambda = 0$ the model degenerates into a random choice model. If q_i is small, subjects react a lot to the payoffs by adjusting their propensity

[10]We are assuming that subjects always solve the easy version of the real effort task successfully.

Table 4.3: Reinforcement learning: Outcome from maximum likelihood estimation

Treatment	λ	**Average** q_i	**Function value**
25-25	0.0087	0.5826	-955.2651
75-75	0.0051	0.8348	-624.4631
25-75	0.0073	0.8183	-726.3666

accordingly while the influence of the previous propensity is limited.

In order to find out if behavior in our experiment can be explained by reinforcement learning, we estimate the parameters q_i and λ for each treatment separately using the maximum likelihood method. The parameters q_i and λ are chosen such that they maximize the following log-likelihood function: $\sum_i \sum_k \sum_t x \log(p_{ikt})$. Variation in subjects' choices over the rounds identifies the parameters. The results are displayed in Table 4.3. Note that we estimate one parameter value for λ in each treatment, but subject-specific discount factors q_i. The estimated discount factors q_i are rather large in all treatments. Thus, the propensity of a strategy highly depends on last round's propensity, while the earned payoff has only a minor impact. Interestingly, the discount factor is smaller in the 25-25 treatment than in the other treatments. When the boosters and the option not to purchase are relatively inefficient, subjects place a higher weight on payoffs and take the past propensity less into account. The relatively small q_is in the 25-25 treatment suggest that subjects switch more frequently between purchasing and not purchasing in this treatment than in the other treatments as could already be seen from Figure 4.8. Comparing our results to the fitted values from other experiments, we find that our qs in the 75-75 and in the 25-75 treatments come close to the one estimated by Nagel and Tang (1998) which is 0.85. The discount factors q in Cabrales et al. (2000) lie in a range between 0.45 and 0.80 such that also our parameter value for the 25-25 treatment does not stick out.

The distribution of the discount factors q_i is illustrated in Figure 4.9. In the 25-25 treatment (top panel), last round's propensity has either a very small or a very large impact on this round's propensity. Consequently, for some subjects the payoff from playing a strategy in the previous round has a large impact on the current propensities, while for other subjects the payoff almost does not affect their current propensities. Hence, behavior differs a lot across subjects within the 25-25 treatment. In the 75-75 treatment, the influence of last round's propensity on this round's propensity is large and thus the earned payoff plays a minor role for most subjects. In the 25-75 treatment (bottom panel), there are more subjects with medium sized discount factors q compared to the other treatments, but again most subjects' propensities depend mainly on last round's propensity and only little on the earned profits. Thus, in the treatments in which boosters work with a probability of 75%, propensities are very persistent, while in the 25-25 treatment there are more subjects who react to payoffs.

Focusing on the probabilities with which each strategy is chosen, it becomes clear that propensities do not have a strong influence on the purchase decisions: in all treatments the parameter λ is very small. Hence, subjects' behavior comes relatively close to random choices. With a value of $\lambda = 0.018$ Nagel and Tang (1998) also estimate a relatively small parameter value, but our values are even smaller. Hence, behavior in our experiment comes closer to random decision making. Cabrales et al. (2000) do not use the exponentialized version of the reinforcement model and thus do not estimate the parameter λ.

The function value describes the goodness-of-fit of the models we are estimating. Comparing the function values across treatments, we observe that the model fits the data best in the 75-75 treatment and worst in the 25-25 treatment. Especially the latter result is not surprising as we have already found out that subjects experiment a lot in this treatment such that their behavior is difficult to predict. As the functional values depend on experiment-specific parameters, it is not useful to compare them to other experiments. They only permit a comparison within our experimental setting.

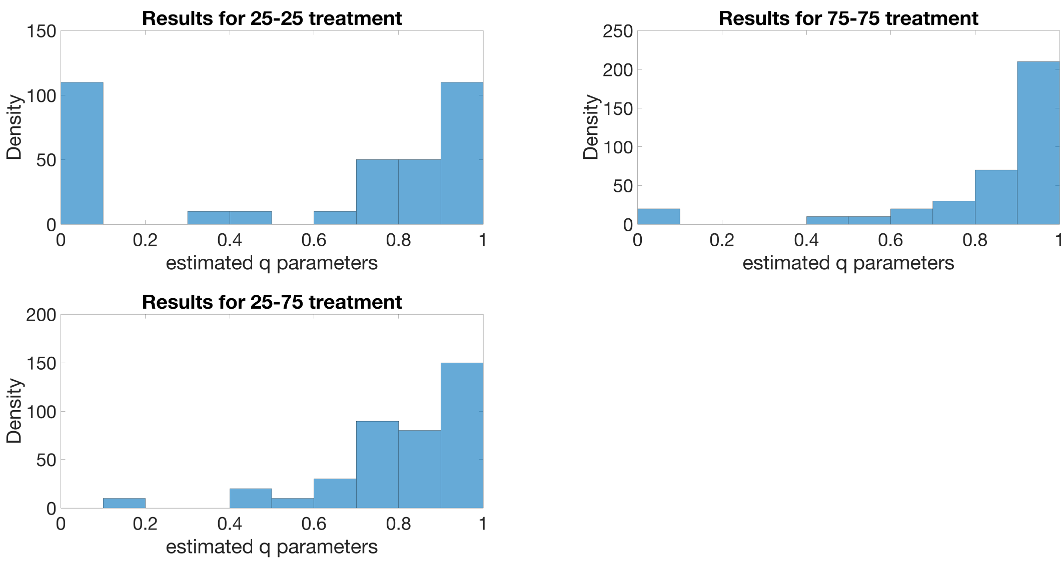


Figure 4.9: Distribution of discount factor q_i

Patterns in buying decisions

To what extend are experienced decision makers attracted to boosters with positive anecdotes? Let us therefore consider Scenario 8 and 10 again. In Scenario 8 two equally priced cheap boosters with different anecdotes are available and in Scenario 10, 20 boosters are offered of which some are cheap and have a positive anecdote. As Table 4.4 shows, on average our subjects are ceteris paribus still more likely to purchase a positive-anecdote booster than a negative-anecdote one in the second half of the experiment. However, the proportion of positive-anecdote boosters purchased is lower than in part 1 (where it was 79%), and in the 25-25 treatment we cannot reject the null hypothesis that subjects pick either type of booster at random.[11] Thus, anecdotes are less important when subjects become more experienced.

Table 4.4: Acquisitions of pos.- versus neg.-anecdote boosters in Scenarios 8 and 10

Treatment	**25-25**		**75-75**		**25-75**	
Scenario	10	8	10	8	10	8
Pos. anecdote boosters	57.1%	55.7%	70.0%	75.0%	59.0%	62.5%
p-value	0.30		0.02**		0.27	

In part 1, subjects were strongly inclined towards low-price boosters. This does not change over the course of the experiment. Table 4.5 summarizes the relevant data for Scenarios 10 (20 boosters), 5 (two boosters with positive anecdotes, different prices) and 6 (two boosters with negative anecdotes, different prices).

What about the relative frequency of S(1)-like behavior? When the prices of positive-anecdote boosters are not systematically higher than those of boosters with negative anecdotes (Scenario 10) 39% of all purchases are precisely and 46% roughly consistent with the S(1)-procedure, compared to 67% in part 1. When positive-anecdote boosters are more expensive – Scenarios 9 and 7 – the proportions of purchases compatible with

Table 4.5: Acquisitions of low-price versus high-price boosters in Scenarios 10, 5 and 6

Treatment	**25-25**			**75-75**			**25-75**		
Scenario	10	5	6	10	5	6	10	5	6
Cheap b.	82.5%	94.7%	83.3%	100%	100%	100%	94.2%	97.9%	96.2%
p-value		$< 0.01^{***}$			$< 0.01^{***}$			$< 0.01^{***}$	

[11]For these one-sample Fisher Pitman permutation tests we pool the data from Scenarios 8 and 10 since the figures within treatments are extremely similar.

S(1) are 14% in Scenario 9 (17% are roughly compatible) and 20% in Scenario 7, compared to 0% (20%) and 33% in part 1. Thus, in relative terms the empirical relevance of the S(1)-procedure across the various scenarios remains similar to what we saw in part 1 – for example the S(1) prediction performs best when boosters associated with positive anecdotes are not more expensive than boosters associated with negative anecdotes – and in absolute terms we find that subjects overall are somewhat less likely to follow the S(1) heuristic when they choose to buy.

4.6 Conclusion

Demand for homeopathy or certain cosmetic products is puzzlingly high in spite of the products' questionable effectiveness. An explanation for the existence of markets for these dubious products may be difficult to construct within the rational choice framework. Spiegler (2006) seeks to explain this phenomenon with a model in which rational sellers face boundedly rational buyers. In his model, consumers' trust in anecdotes is the reason for the existence of markets for useless products: before purchasing an unfamiliar product consumers gather anecdotes about the available costly products and about the option not to purchase. These anecdotes inform the consumer whether the option has previously been successful. If a product has a positive anecdote, consumers are convinced that it will always be successful. Thus, consumers purchase the cheapest product out of those with positive anecdotes. Whenever costly products have received positive anecdotes while the anecdote of the option not to purchase is negative, consumers will demand a costly, but useless product.

Inspired by Spiegler's paper on 'The Market for Quacks', this paper investigates purchasing behavior in situations with very little information. We analyze experimentally whether his behavioral assumption actually describes the decision making behavior of people in situations with little information. Do people always select the cheapest option with a positive anecdote? We seek to find out if other heuristics may be more important, i.e. whether decision makers rather focus on prices or on anecdotes alone. Apart from analyzing behavior in a one shot 'Market for Quacks' setting, we investigate the behavior of more experienced subjects, asking whether the same heuristics are applied when subjects become more familiar with a situation. Confronting subjects with similar, but different situations we ask if experience from one market situation is transferred to similar new market situations.

Seeking to create a situation in which subjects have a certain willingness to pay for products of doubtful efficacy, we design a real effort task that can either be easy or difficult. When it is easy, subjects are usually able to solve the task, whereas the difficult version is almost insolvable. Confronted with the difficult version, subjects are offered costly products which may convert the difficult version into the easy one. They can either choose one of the costly products or decide not to purchase. Both options may transform the difficult version of the real effort task into the easy version. We inform them which of the available products would have simplified the difficult version of the real effort task if they had been confronted with it earlier. Thus, the pieces of information available to subjects ensure comparability to Spiegler's setting: subjects know the prices of the products and we provide them with information about the hypothetical past performance of the products. This piece of information resembles Spiegler's anecdotes. Apart from that, subjects receive no further information. The experiment lasts for at least 60 independent rounds. In each round, the purchase situation changes: the number, the price, the anecdote and the identity of the products are altered. This feature ensures that subjects become familiar with the situation in general, but do not learn more about a specific product.

At the beginning of the experiment we inform subjects only about the first round in which all subjects are confronted with a difficult real effort task (part 1) to capture the static setting of the model. Then, subjects receive the information that 60 additional rounds will be played (part 2). In the second part, our subjects are split into three between-subjects treatments. In the first treatment, both the costly products and the costless option have an efficacy of 25%. In the second treatment, both types of options are again identical, but with a higher efficacy of 75%. In the third treatment, the costly products are superior to the costless option of not purchasing: they transform the difficult version into the easy version in 75% of the cases, whereas the success rate of the costless option lies at 25%. The anecdotes of the boosters are overall representative of the treatment-specific success probabilities (25% or 75%). With these three treatments we seek to find out if behavior changes depending on the success of the chosen options. Furthermore, we want to know if the buying propensity differs across treatments and converges to 0 or 1 depending on the treatment.

In part 1, we find that anecdotes are important for the purchase decisions and that subjects are very price sensitive. Although this observation appears to stand in line with the S(1)-procedure, it only holds when products with positive anecdotes are cheap. If all available products with positive anecdotes are expensive, the observed behavior does not support the S(1)-procedure. In these situations, subjects prefer not to buy or to purchase cheap products with a negative anecdote over S(1). When confronted with only one costly product, the buying propensity is highest when the product is cheap and has a positive anecdote.

In part 2, we observe that anecdotes lose their importance in all treatments, but subjects remain very price sensitive. Additionally, we are surprised to find that buying rates do not converge to 0 in the first two treatments, nor do they converge to 1 in the treatment in which costly products are superior to the costless option. Thus, subjects do not learn to avoid the costly products in the treatments with a costless option as effective as the costly products. Similarly, they do not seek the costly, more effective products in the third treatment when they become more experienced. Although buying rates differ across treatments – the buying rate is lowest when both products and the costless option succeed with a probability of 75% and highest when products are superior to the costless option – the applied decision making heuristics do not vary systematically across treatments.

Focusing on the decision whether or not to purchase while abstracting from specific patterns in the purchase decisions we find that reinforcement learning cannot explain the observed behavior thoroughly. Subjects often do not adjust the propensity to choose a strategy according to its payoff. Furthermore, they tend to switch frequently between purchasing and not purchasing.

One motivation for analyzing the behavior in situations with very limited information was to find out whether assuming that decision makers behave according to the S(1)-procedure is justified. Taking the observations from part 1 and part 2 of the experiment together, we can conclude that the support for the S(1)-procedure in our data is rather mixed: it is only applied when cheap products with positive anecdotes are available. Even if there are low priced options with positive anecdotes, decisions makers do not always behave according to the S(1)-procedure. Thus, the assumption does not fully capture the actual behavior in markets with limited information.

Surprisingly many experienced subjects purchase in the 75-75 treatment and in the 25-25 treatments, whereas notably few purchase in the 25-75 treatment. This observation demonstrates that learning in an environment with limited information is difficult even though the setting in the experiment is highly structured and thus potentially less demanding than naturally occurring situations. Therefore, applying a behavioral model to analyze the decisions in a market with limited information is justified as decision makers struggle to behave rationally.

In the experiment subjects' price-sensitivity dominates all other heuristics. Hence, it appears as though markets for expensive quacks cannot persist. Yet outside of the laboratory, actively managed funds charge considerably higher fees than index funds although their performance does not necessarily justify the price difference. If investors selected their fund according to the S(1)-procedure, the price difference could be explained, but our experiment has shown that this heuristic is not popular for high-priced products. Instead, it may be the trust in experts which decreases the price-sensitivity and generates demand for actively managed funds.

Overall, there is evidence of persisting demand in markets for worthless goods even though people do not always stick to the S(1)-procedure when making purchase decisions. To further analyze the behavior of people in situations with very little information, we plan to run an additional laboratory experiment: in this experiment, real subjects will take the role of sellers. Knowing both the success probability of the costless option for each buyer and the effectiveness of their own products, they need to determine their price. We are curious to find out if purchasing and learning behavior differ when sellers are human. Will buyers apply different heuristics? Potentially they become less price-sensitive when they know that they interact with a human subject. In addition, we want to investigate how the market sides will influence each other. Furthermore, we ask if the success probabilities of the costless options affect the prices set by human sellers.

Appendix

Instructions

Instructions for Part 1

Welcome to the experiment and thank you for participating. In this experiment you can earn points which will be converted into money at the end of the experiment. The conversion rate is 20 points = € 1. You will be paid privately and in cash.
Today's experiment consists of two parts. This text explains the first part. As soon as all participants have completed the first part, the experiment will be paused and you will receive further information about the second part. It is your task to find and hit hidden

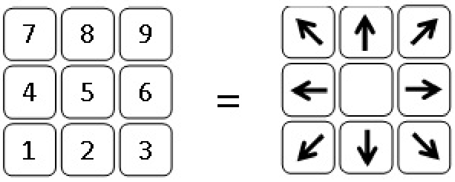
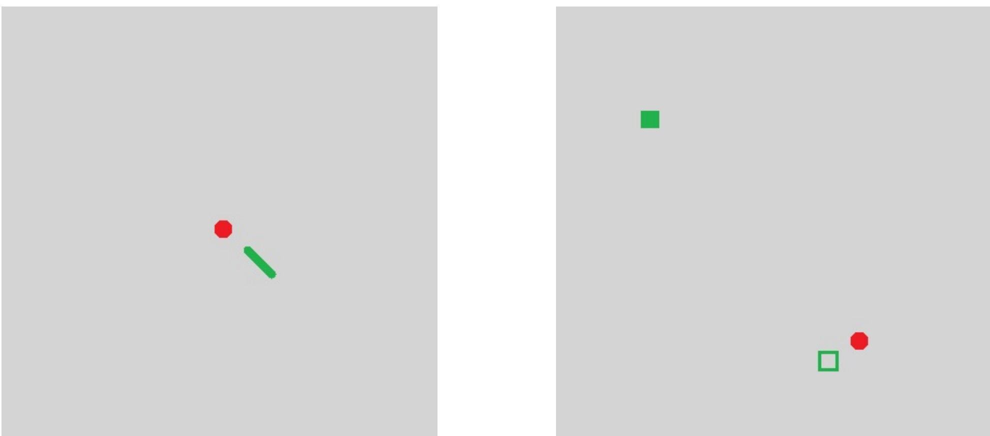


Figure 4.10: Steering the cursor

targets on your screen. In order to search the screen you can move the cursor using with NumPad of your keyboard (see Figure 4.10).
A small green arrow helps you with your task: it points into the direction of the hidden target that needs to be found. In the depicted case (see Figure 4.11) the target is in the bottom right corner. As soon as the cursor is close to the hidden target, the outline becomes visible. This facilitates hitting the target. If the target is hit, it will be shown as green square. You start the first part with an endowment of 100 points. You have 20



(a) Target in bottom right corner **(b)** One found, second within reach

Figure 4.11: Real effort task

seconds to hit all hidden targets. If you succeed, you will receive 100 points in addition.

If you do not succeed, you do not receive any additional points.
It is your task to hit either three or seven targets within a time limit of 20 seconds. Which version of the task you will be confronted with has already been determined and cannot be influenced by your decisions.
You will play five practice rounds before the first part begins in order to prepare you for your task. These practice rounds will include rounds with three hidden targets and rounds with seven hidden targets. Your performance in the practice rounds does not influence your earnings.
If you are confronted with a difficult task (7 targets) after completing the practice rounds, you will have the possibility to purchase a "booster". The purchase of a booster potentially reduces the number of the hidden targets from 7 to 3, but it can also fail. Figure 4.12 depicts a list of available boosters. In the depicted case four boosters are available. In

Figure 4.12: Boosters

an actual situation, there may be more or fewer. The boosters differ in their prices (the depicted prices are random examples). As already mentioned, you start the first part with an endowment of 100 points. You can use these points to purchase a booster. You can only purchase one booster, even though your endowment might allow you to buy several. The category "Successful?" is also included in the booster-table. Important: This does **not** inform you if the respective booster **will be successful** in part 1. Instead, the entry underneath "Successful?" informs you if the booster **would have reduced** the number of targets to three **in the last practice round** (with seven targets). In the depicted table, this would have been the case for Booster c and Booster d, but not for Booster a and b.
Of course, you can also decide not to purchase a booster by choosing the option "Proceed without booster". This option is costless. Potentially the number of targets will be reduced as well, perhaps it stays constant.
In order to get a better overview of the available boosters before making a decision, you can sort the boosters according to the names, their success and their prices. Therefore you

need to click on the bold headlines of the columns. A little triangle next to the category shows whether the boosters are sorted in ascending or descending order. In the depicted case, boosters are sorted by their names in alphabetical order.
If you are confronted with the easy task (3 targets) in part 1, you will not be offered any boosters.
Your earnings in part 1 consist of the initial 100 points plus additional 100 points if you hit all (3 or 7) targets within the time limit minus the points you may have spent on a booster.
You will be informed about the course of the experiment as soon as all participants have completed the first part.

Instructions for Part 2

Having completed the first part of the experiment, you will play at least 60 rounds of the already known game. Please play also the 61^{st}, 62^{nd},... round, if these rounds are displayed. The experiment ends as soon as the last participant has finished the 60^{th} round. The rules in part 2 correspond to the rules in part 1. Please note the following hints:

1. Now, the category "Successful?" informs you if the booster would have been successful in the previous round. If you receive the task to hit seven targets in the first round and the computer offers you boosters, then "successful" boosters are those that would have been successful in the just completed part 1. In round 2, the entry in the category "Successful?" refers to the effectiveness of the displayed boosters in round 1 etc. It does not matter if 7 or only 3 targets had to be hit in the previous round; the entry shows if the respective booster would have reduced the number of targets.

2. You will be offered new boosters in each round with seven targets. For example, a "Booster a" in round 13 and a "Booster a" in round 14 do not have the same identity. It is thus not useful to keep track of the performance of boosters across rounds.

3. The characteristics of the boosters as well as the parameters of the rounds (e.g. number and position of targets) have been pre-determined. Thus, the software does not react to your purchase decisions or to your performance in hitting the targets. You can therefore consider each round separately.

Your payment for the entire experiment is determined as follows: as already mentioned, part 1 is payment-relevant. In addition you receive one payment from part 2. Therefore **one round** of the (at least) 60 rounds from part 2 is **randomly selected**. Your payment from part 2 equals your payment in the selected round. The same conversion rate as in part 1 applies, i.e. you receive €1 for 20 points.

Bibliography

Acebrón, Laurentino Bello and Domingo Calvo Dopico (2000). “The importance of intrinsic and extrinsic cues to expected and experienced quality: an empirical application for beef.” *Food Quality and Preference* 11.3, pp. 229–238.

Anbarci, Nejat and Nick Feltovich (2017). “Pricing in Competitive Search Markets: The Roles of Price Information and Fairness Perceptions.” *Management Science*.

Anbarcı, Nejat, Nick Feltovich, and Mehmet Y Gürdal (2015). “Lying about the price? Ultimatum bargaining with messages and imperfectly observed offers.” *Journal of Economic Behavior & Organization* 116, pp. 346–360.

Anderson, Christopher M (2012). “Ambiguity aversion in multi-armed bandit problems.” *Theory and decision* 72.1, pp. 15–33.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer, and Matthias Sutter (2013). “What drives taxi drivers? A field experiment on fraud in a market for credence goods.” *The Review of Economic Studies* 80.3, pp. 876–891.

Banks, Jeffrey, Mark Olson, and David Porter (1997). “An experimental analysis of the bandit problem.” *Economic Theory* 10.1, pp. 55–77.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964). “Measuring utility by a single-response sequential method.” *Behavioral Science* 9.3, pp. 226–232.

Bergemann, Dirk and Juuso Valimaki (2006). “Bandit problems.”

Blume, Andreas and Andreas Ortmann (2007). “The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria.” *Journal of Economic theory* 132.1, pp. 274–290.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch (2014). “hroot: Hamburg registration and organization online tool.” *European Economic Review* 71, pp. 117–120.

Burdett, Kenneth and Kenneth L Judd (1983). “Equilibrium price dispersion.” *Econometrica: Journal of the Econometric Society*, pp. 955–969.

Burdett, Kenneth, Shouyong Shi, and Randall Wright (2001). “Pricing and matching with frictions.” *Journal of Political Economy* 109.5, pp. 1060–1085.

Cabrales, Antonio, Walter Garcıa-Fontes, and Massimo Motta (2000). “Risk dominance selects the leader: An experimental analysis.” *International Journal of Industrial Organization* 18.1, pp. 137–162.

Cason, Timothy N and Shakun Datta (2006). “An experimental study of price dispersion in an optimal search model with advertising.” *International Journal of Industrial Organization* 24.3, pp. 639–665.

Cason, Timothy N and Charles Noussair (2007). “A market with frictions in the matching process: an experimental study.” *International Economic Review* 48.2, pp. 665–691.

Cason, Timothy N, Roman M Sheremeta, and Jingjing Zhang (2012). “Communication and efficiency in competitive coordination games.” *Games and Economic Behavior* 76.1, pp. 26–43.

Charness, Gary and Dan Levin (2005). “When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect.” *The American Economic Review* 95.4, pp. 1300–1309.

Chen, Yan and Fang-Fang Tang (1998). “Learning and incentive-compatible mechanisms for public goods provision: An experimental study.” *Journal of Political Economy* 106.3, pp. 633–662.

Cooper, Russell, Douglas V DeJong, Robert Forsythe, and Thomas W Ross (1989). “Communication in the battle of the sexes game: some experimental results.” *The RAND Journal of Economics*, pp. 568–587.

— (1992). “Communication in coordination games.” *The Quarterly Journal of Economics*, pp. 739–771.

Corazzini, Luca, Sebastian Kube, Michel André Maréchal, and Antonio Nicolo (2014). “Elections and deceptions: an experimental study on the behavioral effects of democracy.” *American Journal of Political Science* 58.3, pp. 579–592.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang (2007). “Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness.” *Economic Theory* 33.1, pp. 67–80.

Darby, Michael R and Edi Karni (1973). “Free competition and the optimal amount of fraud.” *The Journal of law and economics* 16.1, pp. 67–88.

Dulleck, Uwe and Rudolf Kerschbamer (2006). “On doctors, mechanics, and computer specialists: The economics of credence goods.” *Journal of Economic literature* 44.1, pp. 5–42.

— (2009). “Experts vs. discounters: Consumer free-riding and experts withholding advice in markets for credence goods.” *International Journal of Industrial Organization* 27.1, pp. 15–23.

Dulleck, Uwe, Rudolf Kerschbamer, and Matthias Sutter (2011). “The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition.” *The American Economic Review* 101.2, pp. 526–555.

Dulleck, Uwe, David Johnston, Rudolf Kerschbamer, Matthias Sutter, et al. (2012). *The Good, the Bad and the Naive: Do fair prices signal good types or do they induce good behaviour?* Tech. rep. National Centre for Econometric Research.

Eckel, Catherine C and Philip J Grossman (2008). "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior & Organization* 68.1, pp. 1–17.

Emons, Winand (1997). "Credence goods and fraudulent experts." *The RAND Journal of Economics*, pp. 107–119.

Erev, Ido and Alvin E. Roth (1998). "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria." *The American Economic Review*, pp. 848–881.

Fischbacher, Urs (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics* 10.2, pp. 171–178.

Gehrig, Thomas, Werner Güth, Vittoria Levati, Rene Levinsky, Axel Ockenfels, Tobias Uske, and Torsten Weiland (2007). "Buying a pig in a poke: An experimental study of unconditional veto power." *Journal of Economic Psychology* 28.6, pp. 692–703.

Greiner, Ben (2004). "The online recruitment system ORSEE 2.0." *A Guide for the Organization of Experiments in Economics.*

Grether, David M. (1980). "Bayes rule as a descriptive model: The representativeness heuristic." *The Quarterly Journal of Economics* 95.3, pp. 537–557.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982). "An experimental analysis of ultimatum bargaining." *Journal of economic behavior & organization* 3.4, pp. 367–388.

Helland, Leif, Espen R Moen, and Edgar Preugschat (2017). "Information and coordination frictions in experimental posted offer markets." *Journal of Economic Theory* 167, pp. 53–74.

Holt, Charles A. and Angela M. Smith (2009). "An update on Bayesian updating." *Journal of Economic Behavior & Organization* 69.2, pp. 125–134.

Huck, Steffen, Gabriele K Lünser, and Jean-Robert Tyran (2012). "Competition fosters trust." *Games and Economic Behavior* 76.1, pp. 195–209.

— (2016). "Price competition and reputation in markets for experience goods: An experimental study." *The RAND Journal of Economics* 47.1, pp. 99–117.

Jacoby, Jacob, Jerry C Olson, and Rafael A Haddock (1971). "Price, brand name, and product composition characteristics as determinants of perceived quality." *Journal of Applied Psychology* 55.6, p. 570.

Kriss, Peter H, Rosemarie Nagel, and Roberto A Weber (2013). "Implicit vs. explicit deception in ultimatum games with incomplete information." *Journal of Economic Behavior & Organization* 93, pp. 337–346.

Lester, Benjamin (2011). “Information and prices with capacity constraints.” *The American Economic Review* 101.4, pp. 1591–1600.

Lundquist, Tobias, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson (2009). “The aversion to lying.” *Journal of Economic Behavior & Organization* 70.1, pp. 81–92.

Mimra, Wanda, Alexander Rasch, and Christian Waibel (2016a). “Price competition and reputation in credence goods markets: Experimental evidence.” *Games and Economic Behavior* 100, pp. 337–352.

— (2016b). “Second opinions in markets for expert services: Experimental evidence.” *Journal of Economic Behavior & Organization* 131, pp. 106–125.

Mitzkewitz, Michael and Rosemarie Nagel (1993). “Experimental results on ultimatum games with incomplete information.” *International Journal of Game Theory* 22.2, pp. 171–198.

Morgan, John, Henrik Orzen, and Martin Sefton (2006). “An experimental study of price dispersion.” *Games and Economic Behavior* 54.1, pp. 134–158.

Nagel, Rosemarie and Fang-Fang Tang (1998). “Experimental results on the centipede game in normal form: an investigation on learning.” *Journal of Mathematical psychology* 42.2, pp. 356–384.

Norton, Douglas A and R Mark Isaac (2012). “Experts with a conflict of interest: a source of ambiguity?” *Experimental economics* 15.2, pp. 260–277.

Osborne, Martin J. and Ariel Rubinstein (1998). “Games with procedurally rational players.” *The American Economic Review*, pp. 834–847.

Pesendorfer, Wolfgang and Asher Wolinsky (2003). “Second opinions and price competition: Inefficiency in the market for expert advice.” *The Review of Economic Studies* 70.2, pp. 417–437.

Rabin, Matthew (2002). “Inference by believers in the law of small numbers.” *The Quarterly Journal of Economics* 117.3, pp. 775–816.

Roth, Alvin E. and Ido Erev (1995). “Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term.” *Games and Economic Behavior* 8.1, pp. 164–212.

Schneider, Henry S (2012). “Agency problems and reputation in expert services: Evidence from auto repair.” *The Journal of Industrial Economics* 60.3, pp. 406–433.

Stahl, Dale O (1989). “Oligopolistic pricing with sequential consumer search.” *The American Economic Review*, pp. 700–712.

Tversky, Amos and Daniel Kahneman (1971). “Belief in the law of small numbers.” *Psychological Bulletin* 76.2, pp. 105–110.

Van Huyck, John B, Raymond C Battalio, and Richard O Beil (1993). “Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication.” *Games and Economic Behavior* 5.3, pp. 485–504.

Varian, Hal R (1980). “A model of sales.” *The American Economic Review* 70.4, pp. 651–659.

Curriculum Vitae

2011-2017	Graduate Studies in Economics
	University of Mannheim, Germany
	Center for Doctoral Studies in Economics
08/2017	Ph.D. in Economics
2010-2011	Master Studies in Economics
	Universitat Pompeu Fabra, Spain
	Barcelona Graduate School of Economics
07/2011	Master of Research in Economics
2007-2010	Bachelor Studies in Economics
	University of Bayreuth, Germany
08/2010	Bachelor of Science in Economics
06/2007	Abitur
	Altes Gymnasium Flensburg, Germany