Three Essays on Credence Goods: The Impact of Expert and Market Characteristics on the Level of Expert Fraud

Inauguraldissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Universität zu Köln

2013

vorgelegt von

Dipl.-Vw. Christian Georg Waibel

aus

Münster

Referent:	Prof. Achim Wambach, Ph. D.
Korreferentin:	Prof. Dr. Bettina Rockenbach

Tag der Promotion:	22. Januar 2013

To my parents.

Contents

List of Figures

Acknowledgments

First and foremost, I would like to thank my supervisor Achim Wambach for his guidance and continuous advice. The discussions with him about any of my work were always fruitful and encouraging. I would further like to thank his chair and the European Center for Liberal Professions of the University of Cologne for the financial support of the experiment on which chapter four is based. I would like to express my sincere thanks to Bettina Rockenbach for refereeing this thesis and to Oliver Gürtler for chairing the defense board.

I am very grateful to my colleagues at the chair, Nicolas Fugger, Florian Gössl, Vitali Gretschko, Wanda Mimra and Alexander Rasch for numerous serious but also entertaining discussions. I would especially like to thank Wanda Mimra and Alexander Rasch for their continuous support and co-authoring one respectively two of my papers. I am grateful to Wanda Mimra for financially supporting the experiment. I am also very thankful to Ute Büttner and Susanne Ludewig-Greiner for their great assistance in administrative matters and the several (small) talks we had. Special thanks go to Lisa Boxberg, Nicolas Fugger and Marlene Scholz for assisting Alexander Rasch and me in collecting the data for the article on which chapter two is based. Philippe Gillen deserves my deepest gratitude for supporting Wanda Mimra, Alexander Rasch and me in an excellent way of programming the experiment.

I would like to thank the Cologne Graduate School in Management, Economics and Social Sciences (CGS) for financial support during the last three years. This includes a grant for conducting the experiment. I am grateful to my colleagues at the CGS, Kristian Dicke, David Kusterer, Johannes Mans, Dirk Paulsen and Marta Quintussi for a great time in and out of office at CGS.

I would like to express my sincere thanks to Matthias Sutter and his team at the University of Innsbruck. The research stay in Innsbruck was not only a very productive time but also a great opportunity to get to know the leading researchers in the field of credence goods.

Finally, I would like to thank my parents, Luzia and Gerhard, my brother Matthias and Anja for their comprehensive support not only throughout the doctoral program but already during the time before.

1 Introduction

The present thesis analyzes how expert and market characteristics influence an expert's incentive to defraud his customer in a market for credence goods. A credence good is a good for which the customer does not know which quality of the good she needs (Darby and Karni, 1973). Experts are able to identify the quality the customer needs. This information asymmetry between the expert and the customer may give rise to fraudulent behavior.

The analysis of experts' fraud incentives in credence goods markets is highly relevant. Firstly, there are numerous markets in which the good traded has the credence characteristic: medical treatments, where the physician is better informed about the patient's disease than the patient himself (e. g., Gruber and Owings, 1996; Iizuka, 2007); car repairs, where the mechanic knows better than the owner which parts of the car have to be replaced (e. g., Hubbard, 1998; Taylor, 1995; Wolinsky, 1993); taxi rides, where the cab driver has more information on the shortest route to the destination than the customer (Balafoutas et al., 2013); and lawyers' advice, where the lawyer is better informed about the winning prospects of taking a matter to court than the potential claimant (Dulleck et al., 2012). Secondly, the markets under consideration and the observed levels of fraud are large: the market for medical products and services alone makes up about 10% of the GDP in most industrialized countries (OECD, 2011). In the US, about 2.4 trillion USD are spent on health care each year (OECD, 2011), up to 10% of which are estimated to be due to fraud (Federal Bureau of Investigation, 2007). Finally, credence goods markets do not only make up a relevant share of the GDP today but constantly gain importance in a world of increasingly specialized products and services. Customers are less and less able to evaluate the quality they actually need.

There are three different ways in which the expert may exploit his informational advantage: He may provide high quality (at a possibly higher mark-up) although low quality would have been sufficient, i. e., the expert *overtreats* (see, e. g., Alger and Salanié, 2006; Dulleck and Kerschbamer, 2006, 2007, 2009; Dulleck et al., 2009; Emons, 1997, 2001; Hilger, 2011; Richardson, 1999). If the customer cannot observe the quality provided, the expert may charge for the high quality good although he only provides the low quality good, i. e., the expert *overcharges* (see, e. g., Dulleck and Kerschbamer, 2006; Dulleck et al., 2009; Fong, 2005; Pesendorfer and Wolinsky, 2003; Sülzle and Wambach, 2005; Taylor, 1995; Wolinsky, 1993, 1995). If the customer needs high quality but the expert only serves low quality, the expert *undertreats* (see, e. g., Alger and Salanié, 2006; Dulleck and Kerschbamer, 2006, 2007; Hilger, 2011; Richardson, 1999). The customer does not notice whether she was overcharged or overtreated even ex-post in a credence goods market, but the customer is able to verify whether she was undertreated. In this thesis, I focus on the expert's incentive to overcharge. The fourth chapter additionally investigates the expert's incentive to undertreat.

The remainder of this thesis is organized as follows. The second chapter entitled *What Drives Fraud in a Credence Goods Market – Evidence From a Field Study* is based on joint work with Alexander Rasch.[1] The work is motivated by the fact that the existing literature on field studies in credence goods markets has mainly focused on customer characteristics (e. g., Balafoutas et al., 2013). We contribute to the literature by extending the analysis firstly to market characteristics, such as the

[1]I presented the results from this study at the *10th International Organization Conference* (Washington D. C., United States, 2012), at the *Economics Seminar* (University of Augsburg, Germany, 2012), at the *European Economic Association* (Malaga, Spain, 2012), at the *European Association for Research in Industrial Economics* (Rome, Italy, 2012), at the *European Science Association* (Cologne, Germany, 2012) and at the *Applied Microeconomics Seminar* (University of Cologne, Germany, 2012).

degree of competition, and secondly to expert characteristics.[2] On the expert side, we analyze the impact of the expert's financial situation, competence, and reputational concerns on the expert's incentive to overcharge. We resort to the unifying model of Dulleck and Kerschbamer (2006) and extend it in order to make predictions about the influence of these characteristics. We make use of a unique dataset on car repair shops provided by Germany's largest automobile club to test our theoretical predictions. The automobile club provides information on overcharging and the firms' competence. We extend this dataset by collecting the number of garages in a ten-kilometer radius from a garage's location in order to quantify the intensity of competition. Furthermore, we determine a garage's geographical proximity to the next interstate and use it as an indicator for a lower share of repeated business contacts and hence less reputational concerns. Last, we collect data about the firm's financial situation from the Electronic Federal Gazette (*Bundesanzeiger*). We restrict our dataset to corporate garages. The restriction is feasible because only corporate garages have to publish their financial situation. Furthermore, we derive our predictions from a model assuming limited liability. Corporate garages mostly operate under limited liability while non-corporate garages do not. We show that a higher degree of competition lowers the incentive to overcharge. We find that firms facing a critical financial situation are more likely to overcharge. Garages with a high competence are less likely to overcharge than those with a low competence. Our results also indicate that less reputation-oriented car repair shops defraud their customers more often than those with high reputational concerns. These results are in line with our theoretical predictions.

The third chapter entitled *Fraud and Other-Regarding Preferences in a Market for Credence Goods* theoretically analyzes the impact of customers' and experts' other-regarding preferences on the expert's incentive to overcharge.[3] Previous work

[2]To the best of my knowledge, the only field study concerned with expert characteristics is Schneider (2012). He analyzes the impact of reputational concerns on the expert's incentive to defraud.

[3]I presented the results from this study at the *8th World Congress on Health Economics* (Toronto, Canada, 2011), at the *10th Annual International Conference on Health Economics* (Athens, Greece, 2011), at the *Spring Meeting for Young Economists* (Groningen, Netherlands, 2011), at

on experts' behavior has mainly focused on monetary incentives (e. g., Dulleck and Kerschbamer, 2006; Emons, 1997, 2001; Wolinsky, 1993, 1995). However, experimental results show that experts like physicians base their decisions to a "considerable extent" on norms like honesty (Dulleck et al., 2011). Even in a situation where theory would predict that experts always defraud their customers, only half of them actually do so. We therefore contribute to the literature by analyzing how an expert's preference for honesty and a customer's preference for trust influence the expert's incentive to overcharge. Physicians face *conscience costs* in our model when they overcharge their patients. Patients face *trust costs* when being charged a major treatment because they anticipate that they may have been defrauded. We study the impact of these preferences on the level of overcharging when patients can consult a second physician. Contrary to intuition, we show that the level of fraud does not necessarily decrease in equilibrium but may even rise compared to players without other-regarding preferences. Welfare increases with experts' honesty but reacts ambiguously to increased patients' trust costs.

The fourth chapter entitled *Reputation in Credence Goods Markets—Experimental Evidence* is joint work with Wanda Mimra and Alexander Rasch.[4] Previous experimental work has studied credence goods markets where customers build up a private history with respect to the sufficiency of a treatment and the prices charged (see Dulleck et al., 2011). This work firstly extends the existing literature by analyzing the impact of public histories on the experts' incentive to defraud his customer. In contrast to private histories, public histories imply that customers do not only know their own history but customers observe all customers' histories. The analysis of public histories is motivated by the recent emergence of feedback platforms. A prime example is the *Arztnavigator*[5] where patients pass on their experience with specific physicians, aiming to help other patients compare the quality and the price

the *3rd Annual Conference of the German Society for Health Economics* (Bayreuth, Germany, 2011), and at the *Applied Microeconomics Seminar* (University of Cologne, Germany, 2012).

[4] I presented this work at the *Doctoral Seminar* (University of Innsbruck, Austria, 2012).

[5] The *Arztnavigator* polls the patients with a standardized questionnaire about their last physician visit. The results are cumulated per physician and publicly released (source: `http://www.arztnavigator.de`).

of physician services. The feedback platform thereby gives physicians the opportunity to build up a reputation publicly.[6] The second contribution in this chapter is to analyze how the pricing regime impacts an expert's incentive to defraud his customer if experts can build up reputation. Whereas in the market for medical treatments, legal services, and taxi rides, prices are regulated in most European countries, experts set individual prices in the car repair business. We experimentally investigate how the possibility to build up reputation and how the pricing regime influence an expert's incentive to undertreat and to overcharge his customer. We make use of a 2×2 factorial design and vary the reputation mechanism between private and public histories and the price system between flexible and fixed prices. The results show that the level of undertreatment and the level of overcharging are significantly lower under fixed than under flexible prices. In contrast to intuition, the results provide weak evidence that the level of overcharging is higher under public than under private histories if prices are flexible while the opposite holds when prices are fixed. We find the same pattern for the level of undertreatment but differences are not significant.

The last chapter concludes. It discusses the implications of the studies presented and gives an outlook for future avenues of research.

[6] Note that the patient can ex-post verify whether she is undertreated (or inappropriately treated) because her disease is not cured. This is in contrast to overtreatment where the patient does not take notice of the fraudulent behavior because she is healed.

2 What Drives Fraud in a Credence Goods Market—Evidence From a Field Study

This paper investigates the impact of four key economic variables on an expert firm's incentive to defraud its customers in a credence goods market: the level of competition, the expert firm's financial situation, its competence, and its reputational concerns. We use and complement the dataset of a nationwide field study conducted by the German Automobile Association that regularly checks the reliability of garages in Germany. We focus on corporate garages and find that more intense competition and high competence lower firms' incentive to overcharge. A low concern for reputation and a critical financial situation increase the incentive to overcharge.

2.1 Introduction

In this paper, we analyze the impact of expert and market characteristics on an expert firm's incentive to defraud its customers in a credence goods market. We make use of a field study in the German car repair market in order to identify the drivers of fraudulent behavior. Faulty repairs and fraudulent behavior are major issues in this market. According to a joint survey by the Consumer Federation of America, the National Association of Consumer Agency Administrators, and the North American Consumer Protection Investigators, faulty repairs in the auto repair market rank first among the top ten consumer complaints in 2010. The California Department of Consumer Affairs notes that complaints related to car repairs also grew fastest during the same period. Its Bureau of Auto Repair even shut down some shops of one chain due to overcharging and overtreatment (Consumer Federation of America et al., 2011). These results are in line with earlier studies which also found that fraud related to auto repairs was among the most often observed types of fraudulent behavior.[7] In this paper, we focus on the expert's incentive to overcharge.

Given the role fraud (and overcharging in particular) plays in this market, it is important to better understand the factors that make it easier or harder for experts to exploit their informational advantage at the expense of their customers. In order to analyze experts' overcharging behavior, we make use of the results from a field study in the German car repair market that is carried out on a yearly basis by the German Automobile Association (*Allgemeiner Deutscher Automobil-Club e. V.*, ADAC), Europe's largest automobile club. The ADAC has looked into the reliability of German car repair shops over several years. We are interested in the influence of four key economic variables on expert firms' incentives to defraud their customers: competition, financial status, firm competence, and reputation. By analyzing the impact of these economic variables, our study complements other contributions that have focused on different determinants of fraudulent behavior (see below). In contrast to earlier contributions, we focus on expert rather than customer characteristics. Furthermore, by considering the degree of competition, we account for an important market characteristic.

The automobile club's database contains information on overcharging and the firms' competence. The automobile club recorded overcharging if the number of repairs charged exceeded the number of faults fixed. We extend this database by collecting the number of garages in a ten-kilometer distance from a garage's location

[7] See, e.g., Titus et al. (1995). See also the study by the U.S. Department of Transportation cited in Wolinsky (1993, 1995). A 2002 poll conducted by COMsciences, Inc. for Allstate Insurance Company revealed that there was a general atmosphere of distrust in auto body repair shops among consumers in California: among others, consumers were concerned about cheating and inflated prices (see *Business Wire*, August 12, 2002, Monday: "Survey shows Californians fed up with auto repair fraud; pending legislation threatens to block reform and restrict competition").

in order to quantify the intensity of competition. Furthermore, we determine a garage's geographical proximity to the next interstate and use it as an indicator for a lower share of repeated business contacts and hence less reputational concerns. Last, we collect data about the firm's financial situation. In our main analysis, we focus on corporate garages because only corporate garages have to publish their financial situation. Furthermore, we derive our predictions from a model assuming limited liability. Corporate garages mostly operate under limited liability while non-corporate garages do not.

We show that a higher degree of competition lowers the incentive to overcharge. We find that firms facing a critical financial situation are more likely to overcharge. Garages with a high competence are less likely to overcharge than those with a low competence. Our results also indicate that less reputation-oriented car repair shops defraud their customers more often than those with high reputational concerns. These results are in line with our theoretical predictions.

The seminal theoretical contribution on fraud in the car repair market is Taylor (1995): he studies an expert's incentive to overcharge his customer. The author shows that under short-term contracts, experts will charge all customers for a treatment independent of whether the car is faulty or not. Consequently, all customers whose car is not faulty are overcharged. In contrast to that model, we assume that customers are not committed to a certain expert, i.e., customers can search for a second opinion after receiving the diagnosis. The reason we make use of a model that captures second opinions is based on the way a car repair market functions. We often observe that mechanics first suggest a treatment and then ask for customers' approval before performing the treatment.

There exist only few experimental/empirical studies focusing on the determinants of dishonest behavior in markets for credence goods. Balafoutas et al. (2013) perform a field experiment on credence goods concerning taxi rides in Athens, Greece. The authors focus on the impact of customer characteristics on the expert's incentive to cheat. Their study reveals that if passengers have only poor information about optimal routes, they are taken on longer detours. The authors also point out that a higher (perceived) customer income increases the level of fraud.[8]

A related study to ours is the recent work by Schneider (2012): similar to our paper, he is interested in garages' (dis)honest behavior toward customers. Schneider (2012) analyzes data from a field experiment where he visited garages undercover in order to check whether expert reputation may alleviate the efficiency problems arising from asymmetric information. He finds both pervasive overtreatment and undertreatment but no evidence that reputation helps reduce these problems.[9]

In the present study, we are the first to explore the influence of market characteristics on the level of overcharging in the field. More precisely, we analyze the impact of competition on expert firms' incentive to defraud their customers. In the competition policy debate, the level of competition among car repair shops often comes up as an important issue: for example, in the above-mentioned poll performed by COMsciences, a great majority of participants supported increased competition in auto repair (e.g., through insurance-owned shops) in order to reduce widespread fraud. Interestingly, the aspect of competition in credence goods markets has only been studied from a theoretical perspective or in the laboratory. Furthermore, we investigate the effect of essential expert characteristics which are not accounted for in Schneider (2012). The experts' financial situation as well as their competence plays a crucial role in the experts' decision on whether to overcharge the customer. Again, the 2002 COMsciences poll revealed that an "overwhelming majority (74%) [of consumers] fear they are often cheated by auto body repair shops that do poor quality work." Moreover, we provide theoretical predictions on these effects from an extension of the unifying model in Dulleck and Kerschbamer (2006). Our study

[8] Dulleck et al. (2011) provide the first experimental study on credence goods. Their main focus is on the role of liability and verifiability in credence goods markets and consider reputation as an extension. They show that neither competition nor reputation decreases the experts' incentive to overcharge. In their empirical study on restaurant hygiene, Jin and Leslie (2009) find that chain-affiliated restaurants have a better hygiene than independent restaurants. This is due to the reputational effects caused by the affiliation.

[9] He also shows that there is a positive relationship between the level of capacity available at a garage at the time of the visit and the probability of a repair recommendation. Moreover, there is a repeat-business effect for the diagnosis fee.

is also based on a larger dataset than Schneider (2012) which allows us to draw more comprehensive conclusions on the underlying causes for fraudulent behavior. Moreover, whereas Schneider (2012) pools data from two different studies, we can revert to data from a single study.

The remainder of the paper is organized as follows: in the next section, we derive our hypotheses from the theoretical literature on credence goods. We describe the dataset in section 2.3. In section 2.4, we present our results and compare them to the theoretical predictions. We check the robustness of our results in section 2.5. The last section concludes and discusses implications for other credence goods markets.

2.2 Theoretical Predictions

For the theoretical analysis, we make use of the model by Dulleck and Kerschbamer (2006) to derive our hypotheses. We present the basic underlying incentives which help explain firms' incentive to overcharge.[10]

Consider the following car repair market. There is a mass one of homogeneous customers (car owners) who all either face a major or a minor problem which occurs with an ex-ante probability of h and $1 - h$, respectively. The problem can be fixed through a major or minor treatment[11], respectively. Customers do not know which type of treatment they require. On the other hand, there is a number of liable expert firms (garages) n (with $n \geq 2$) which are able to diagnose the treatment needed. Liability implies that experts cannot provide a minor treatment to customers facing a major problem, i.e., experts cannot undertreat their customers. Experts set treatment prices and incur costs for providing a treatment. The minor treatment induces costs c_L that are lower than for the major treatment c_H. Experts set a price

[10] An extensive review of the theoretical literature and a unifying model are given in Dulleck and Kerschbamer (2006).

[11] We apply the notion of minor and major treatment used in the credence goods literature. In the real-life market we analyze, the minor treatment corresponds to performing no treatment while the major treatment corresponds to performing a treatment.

p_L for the minor treatment and a price p_H for the major treatment.[12] Assuming that the customer cannot verify the type of treatment, experts have an incentive to overcharge customers with a minor problem by providing a minor treatment but charging for a major treatment. Customers get utility v if their problem is fixed and zero otherwise. They incur search costs of d (due to time and effort) per expert consulted independently of whether they accept the expert's recommendation. We assume that these costs are not too high ($d < (c_H - c_L)(1 - h)$), i.e., economies of scope are sufficiently low. This appears to be a reasonable assumption for inspections in the car repair market which follow a well-established routine. We also assume that it is always (i.e., even ex post) efficient that any customer with a problem is treated which means that $v - c_H - d > 0$ holds.[13] Note that—compatible with the car repair market—we consider the case where a customer is not committed to undergo the treatment recommended by the expert but may decide to spend additional per-visit search costs d on a second, third, etc. opinion instead. Moreover, customers are able to verify whether their problem has been fixed or not.

In this setup, there exists an equilibrium which is characterized as follows:[14] expert firms set prices such that they make a positive profit on minor treatments whereas marginal-cost pricing occurs for the major treatment. Experts always recommend the major treatment if needed but also recommend the major treatment with strictly positive probability x if the customer only needs the minor treatment, i.e., overcharging occurs with strictly positive probability.[15] On the other hand, customers always accept a minor recommendation but visit a second expert with positive probability $1 - y$ if they are recommended the major treatment. On their

[12]We assume that there is a lower bound equal to marginal costs c_H and an upper bound equal to $c_H + d$ for the price of the major treatment. The assumptions map to the car repair market because most car producers enjoin garages on a price range for inspections.

[13]We further assume that customers who are indifferent between visiting an expert and not visiting an expert opt for a visit. Customers who decide for a visit and are indifferent between two or more experts randomize (with equal probability) among them.

[14]See part (i) of *Lemma 6* in Dulleck and Kerschbamer (2006).

[15]See also Pitchik and Schotter (1987, 1993), Wolinsky (1995), Fong (2005), as well as Sülzle and Wambach (2005) for outcomes with overcharging.

second visit, they accept any recommendation with certainty. Moreover, a customer is never undertreated due to the experts' liability.

In such a market, two incentive-compatibility constraints play an important role: an expert firm consulted by a customer with a minor problem finds it more (less) profitable to cheat rather than treat its customers honestly if and only if

$$p_L - c_L < (>) \frac{y + x(1-y)}{1 + x(1-y)} (p_H - c_L) \tag{2.1}$$

The left-hand side gives the profit from honest treatment. Accordingly, the right-hand side represents the gains from recommending the major treatment. Note that in this case, the fraction $1/(1 + x(1-y))$ of customers are on their first visit and accept the high recommendation with probability y. $x(1-y)/(1 + x(1-y))$ customers are on their second visit and accept a high recommendation with certainty.

Similarly, a customer prefers (does not prefer) to seek a second opinion if and only if

$$d < (>) \frac{x(1-h)}{h + x(1-h)} (1-x)(p_H - p_L) \tag{2.2}$$

d represents the additional costs of searching for a second opinion whereas the right-hand side of the inequality gives the expected savings from visiting a second expert firm. Note that with probability $x(1-h)/(h + x(1-h))$, the customer suffers from a minor problem given a major recommendation at the first visit. With probability $1-x$, the second expert honestly recommends the minor treatment. In this case, the customer saves the cost differential $p_H - p_L$ compared to the first recommendation.

Taking this market as a starting point, we use the two inequalities given in (2.1) and (2.2) to motivate our hypotheses. We first look at the relation between competition and overcharging:

Hypothesis 2.1. *As the degree of competition among expert firms intensifies, firms tend to overcharge less.*

We extend the above model by assuming that the customers' search costs d depend on the number of firms n that are located in a customer's neighborhood. The more garages there are in a customer's neighborhood, the lower are the search costs, i.e., $d'(n) < 0$. This is due to the fact that customers have to spend less time and effort searching for suitable experts. Formally, customers' optimal search decision is determined by

$$d(n) < (>) \frac{x(1-h)}{h + x(1-h)}(1-x)(p_H - p_L) \tag{2.3}$$

Ceteris paribus, customers look out for a second opinion at a lower cost as the left-hand side decreases in the number of firms. Consequently, they are more likely to reject a major treatment recommendation. This in turn decreases the firms' incentive to overcharge (see *Lemma 2.1* in *Appendix A*).

Next, we have a closer look at the impact of a lower financial status on overcharging:

Hypothesis 2.2. *An expert firm in a critical financial situation is more likely to overcharge its customers.*

Suppose a firm in the above-described market additionally has to bear fixed costs f in order to run its business and firms differ in their financial assets (low and high). Now, if a firm lacks sufficient financial resources to survive the current period if it does not attract any customer, it does not pay the fixed costs in case it goes bankrupt due to limited liability.[16] As a consequence, it faces lower costs and hence higher profits whenever it recommends the major treatment compared to the firm with the sound financial background. As a result, this firm's optimal recommendation choice then depends on

$$p_L - c_L - f < (>) \frac{y + x(1-y)}{1 + x(1-y)} \left(p_H - c_L - f\right).$$

[16]Note that the assumption of limited liability is satisfied for most of the firms in our dataset.

This means that, all things equal, whenever the financially weak expert firm does not find it profitable to cheat, this is even less the case for the financially strong firm. Hence, the latter has a lower incentive to defraud its customers because it gains more by recommending the minor treatment whenever it is needed (see *Lemma 2.2* in *Appendix A*).

Next, we look at the influence of a firm's competence on its incentive to defraud its customers:

Hypothesis 2.3. *A high-competence expert firm is less likely to overcharge than a low-competence firm.*

Suppose a high-competence firm in our market has lower treatment costs than a low-competence firm. This is captured by a reduction of γ of the initial costs for each treatment which may be due to, e.g., less time-consuming fault detection. Compared to a low competence firm, a firm with high competence only benefits from its better cost situation with certainty if it recommends the minor treatment. If it recommends the major treatment, it may realize the cost advantage only with a probability strictly smaller than one. More precisely, all things equal, the optimal recommendation decision depends on

$$p_L - (c_L - \gamma) < (>) \frac{y + x(1-y)}{1 + x(1-y)} \left(p_H - (c_L - \gamma)\right).$$

As a consequence, the high competence firm faces relatively higher costs and lower profits whenever it recommends the major treatment. Similarly to the above argument in the context of fixed costs, this means that whenever it is not optimal for the low-competence expert firm to cheat, cheating is an even less profitable option for the high-competence firm. As a result, the former has a greater incentive to defraud its customers (see *Lemma 2.3* in *Appendix A*).

Last, let us have a closer look at the relation between reputation and overcharging:

Hypothesis 2.4. *Experts with low reputational concerns are more likely to overcharge than experts with high reputational concerns.*

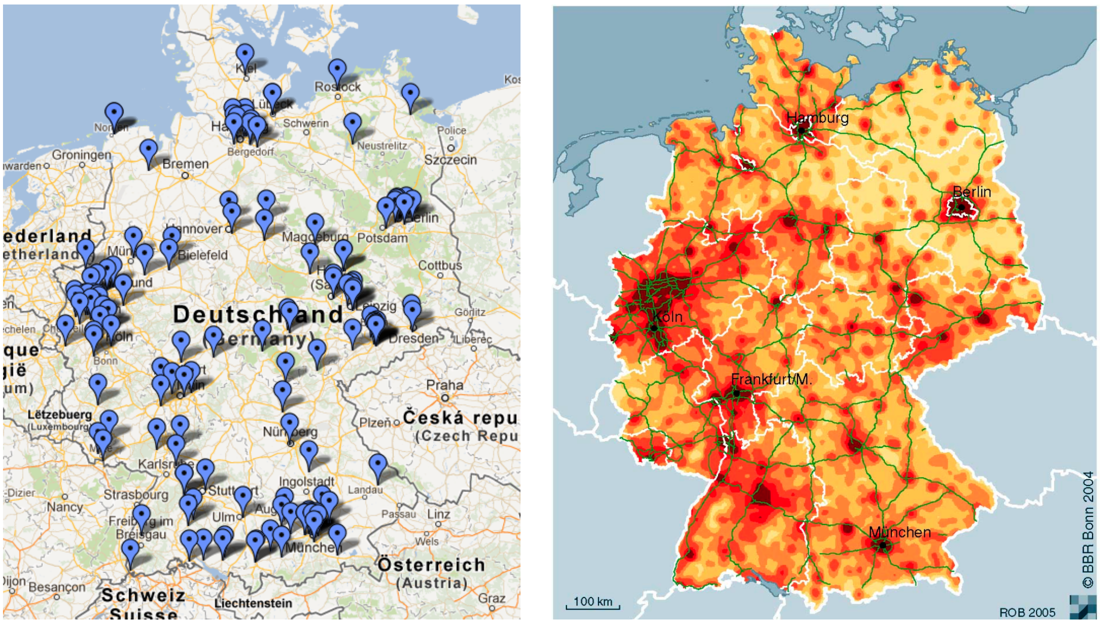
Experts with high reputational concerns face many repeated interactions. Dulleck et al. (2011) show that repeated interaction decreases the incentive to overcharge as experts find it optimal to forgo short-term profits from overcharging because they benefit more from higher profits due to reputation in the future. In line with these findings, Wolinsky (1993) and Park (2005) find that the need to maintain a good reputation decreases the incentive to defraud.

2.3 Data

2.3.1 Sample

We make use of pooled cross-section data from the ADAC's garage tests in the years 2006 and 2008–2010; in 2007, there was no test.[17] The automobile club's dataset provides information on 336 garages. We disregard 39 garages that belong to the same corporate entity because these observations are not independent. 297 observations remain. In our main analysis, we further restrict the sample to 134 corporate enterprises because of data availability and firm characteristics: firstly, only corporate enterprises have to publish data on their financial situation. As we shall see later, a garage's financial situation is an important predictor for the garage's incentive to overcharge. Thus, not considering the financial situation would lead to an omitted variable bias in the estimates. Secondly, the group of corporate garages is a homogeneous subset of all garages while non-corporate garages differ to a greater extent in their properties. The data shows that the variance in competition intensity, competence, and reputational concerns is larger for non-corporate than for corporate garages. Thirdly, we derived our theoretical predictions based on a model that assumes firms to operate under limited liability. This is the case for almost all corporate but not for non-corporate garages. Hence, restricting the dataset to the corporate enterprises seems reasonable. We discuss the case of the unrestricted sample with the 297 independent garages in our robustness section.

[17] See `http://www.adac.de/infotestrat/tests/autohaus-werkstatt/` for details.



(a) Garage locations across Germany (source: Google Maps).

(b) Population density across Germany (source: Federal Institute for Research on Building, Urban Affairs and Spatial Development).

Figure 2.1: Location of garages and population density in Germany.

The location of the 134 corporate garages closely follows the population density within Germany. *Figures 2.1(a)* and *2.1(b)* illustrate this relationship.

The timing of the data collection is as follows:

1. Club members from all over Germany are asked whether they would like to participate in the garage test.

2. The automobile club checks whether the cars fit the test criteria. The cars have to be similar with respect to maintenance-related characteristics (concerning effort and time required): all cars had to be registered during the same time period for the first time, have a gasoline engine (of the most popular performance type), have to be due for the main inspection, and the owners need to present a detailed record of previous inspections.

3. Motor vehicle experts prepare the cars with the same five faults. The faults are the following: the license plate lamp does not work; the air pressure in the spare wheel is too low; the exhaust is loose; the coolant level is low; and the

front-right light is displaced to the very bottom. If any of these faults cannot be implemented, the screen wiper blade on the passenger side is cut down to two centimeters. These potential faults are all listed in any of the car makers' inspection guidelines which means that they should be easily detected.

4. The automobile club sends these cars off to garages located in the vicinity of the car owner's residence. There is a maximum of one vehicle test per garage.

5. Upon completion of the inspections, the automobile club assesses each garage's performance according to a detailed evaluation scheme that also includes issues related to service etc. The results are published in the club's monthly magazine (*ADAC Motorwelt*) and can be readily accessed online. The automobile club gives detailed reports on each garage by exactly listing how many faults were found and fixed and whether only those repairs actually performed were charged.

Our binary dependent variable *overcharging* indicates whether a garage charged for a repair it did not perform. Note that our data only covers parts of the garages' overcharging behavior as we can only determine whether or not a garage charges more repairs than performed. We cannot account for more expensive repairs charged than performed. We consider the number of faults detected by the garage from the automobile club's dataset as an indicator for a garage's competence.

This very basic dataset does not allow us to investigate the impact of the other three key economic variables we are interested in: competition, the firm's financial situation, and its reputational concerns. In order to analyze their influence, we need to complement the automobile club's dataset. This is done in three steps: we (i) introduce a measure for the competitive environment each of the garages does business in, (ii) check for the garages' financial indicators, and (iii) suggest a proxy for reputational concerns (see *Appendix B* for screenshots of the data collection). *Table 2.1* provides an overview over the variables, the proxies, and the respective data sources.

Table 2.1: Overview on variables.

Variable	Proxy	Source
Overcharging	Treatments charged but not performed	ADAC experiment, 2006 & 2008–2010
Competition intensity	# of competitors within 10km is above median	*Gelbe Seiten* from 2011
Financial situation	Negative equity	*Elektronischer Bundesanzeiger*, 2006 & 2008–2010
Competence	# of faults found out of 5	ADAC experiment, 2006 & 2008–2010
Low reputation	Distance to next interstate less than 1500m	Google Maps Distance Calculator, 2010

Ad (i): in order to evaluate the strength of the competition a garage faces, we analyze the number of competitors in a garage's neighborhood. We chose the number of competitors as an indicator for competition over other measures such as the Hirschman-Herfindahl Index (Hirschman, 1964) and the price-cost margin (Boone, 2008) because of data availability. Note that the number of competitors has been used as a proxy for competition by other studies in credence goods markets before (see, e. g., Pike, 2010) but it has not been tested how well the number of competitors actually proxies the competition intensity. Thus, the validity of the proxy we use remains to be shown.

We collect the number of garages that are within a distance of ten kilometers from the garage that is characterized. We consider ten kilometers to be the average distance a potential customer is willing to travel to a competitor.[18] We obtain the data on the number of competitors of every single garage through a request to the publicly available directory of businesses sorted by branches, the German version of yellow pages (*Gelbe Seiten*). *Gelbe Seiten* provides one of the largest phone and

[18] Our results do not change if we take five or 20 kilometers as the radius a customer is willing to travel (see section 2.5 for robustness checks).

address lists of companies in Germany.[19] The great advantage of this database compared to, e. g., *Google Places*, is that the editing process ensures that businesses listed actually exist and fall into the category of car repair shops. We perform a search for "Autowerkstätten" ("car repair shops") within a radius of ten kilometers from the garage's address and count the number of results. Last, we divide the group of garages into those being above the median number of competitors and those below. By dichotomizing competition intensity, we account for the fact that garages' overcharging behavior most likely depends upon whether there are few or many competitors but not on whether there are one or two additional competitors in the nearer neighborhood. Note importantly that our results do not rely on the dichotomization of the variable as shown in the robustness section.

Ad (ii): we extend the automobile club's dataset by adding the garages' financial situation at the beginning and the end of the test year. The financial data is publicly available through the Electronic Federal Gazette for corporate enterprises in Germany (*elektronischer Bundesanzeiger*).[20] According to German corporate law, enterprises are required to publish basic financial information for possible shareholders. In case the balance information was not available by August 2011, we proxied the financial data by using the data from the year before. We divide the garages into those with positive equity and those with negative equity either at the beginning or the end of the year. A firm faces negative equity if its debts exceed its assets. These firms are in a critical financial situation because banks are no longer willing to lend additional money. Firms with a negative equity are not yet bankrupt, though. Bankruptcy is only reached if one of the debts is due and cannot be paid back to the lender. As the amount of a firm's equity is correlated with firm size, we dichotomize the equity variable. Hence, we only capture the firm's financial status without confounding the status with firm size. We chose to use equity as a proxy for a firm's financial situation over other indicators such as profit because equity is not subject to yearly up- and downturns. In particular, equity is invariant with respect

[19]See `http://www.gelbeseiten.de` for details.
[20]See `http://www.bundesanzeiger.de` for details.

to depreciation. Again the drawback of the self-definition of our proxy is that the validity of the proxy has not yet been investigated by other studies.

Ad (iii): we extend the database by adding the garages' distance to the next interstate. We consider this distance as a good proxy for a garage's reputational concerns. Cars that break down on the interstate are usually towed to the next garage.[21] This means that those garages that are located close to an interstate face more one-time interactions. More one-time interactions imply a lower chance of repeat business. As a consequence, they are less concerned when it comes to building up a reputation compared to the garages that are located further away from an interstate. We consider garages that are located less than 1500 meters away from an interstate to be close and all others not to be close to an interstate.[22] We dichotomize the distance to the next interstate because cars are hardly ever towed to a garage that is far away from the interstate. This holds irrespective of whether the garage is ten or 30 kilometers away from the next interstate. We complement the dataset by the garages' exact distances to the next interstate which we calculate using *Google Maps Distance Calculator*. The *Google Maps Distance Calculator* uses *Google*'s geographic database via APIs and enables the user to select two arbitrary points on the map in order to calculate the air-line distance. We take the garage's address as the reference point and the closest point on the next interstate as the second point.[23]

There might be reverse-causality concerns for the relationship between reputational concerns and overcharging as well as the level of competition and overcharging. This is because the choice of a garage's location and thus the distance to the next interstate and the level of competition might not be exogenous to explain overcharging. There are three reasons why we think that a garage's location is in-

[21] The vast majority of the overall number of towings in Germany are conducted by the ADAC. The ADAC always tows to the next garage as their free service for members. Having one's car towed to any other garage is subject to a service fee (see `http://www.adac.de/mitgliedschaft/leistungen/default.aspx`).

[22] Our results are robust if we consider garages less than 1000 meters or less than 2000 meters away from the next interstate as being close to the interstate (see section 2.5).

[23] See `http://www.daftlogic.com/projects-google-maps-distance-calculator.htm` for details.

deed exogenous: firstly, the average age of the garages that overcharged in the test amounts to 20 years (the minimum age to ten years). The garage's overcharging behavior today would have to be correlated to the choice of location twenty years ago if endogeneity concerns were to hold. Hence, a reverse causality does not seem very plausible. Secondly, garages cannot be located anywhere but have to be opened up within a zoned area. Thus, garages are not free to choose a location but are restricted in their choice of location. Thirdly, asking business insiders about where to open new garages provides a clear message: maximizing customer visits is the main goal.[24] These three reasons strengthen our argument that the location is not chosen with respect to the type of interaction (i. e., repeated or one-time) or the number of competitors.

Reverse causality between the incentive to overcharge and a garage's financial situation might exist. As overcharging influences the firm's financial situation, we might encounter endogeneity when considering the equity at the end of the year. Note, however, that overcharging increases equity compared to an honest repair. Consequently, if there was reverse causality between overcharging and a firm's equity, we underestimate the effect of the financial situation on the probability of overcharging. Thus, reverse causality with respect to the financial situation would weaken our results.

2.3.2 Descriptives

After restricting the dataset, the dataset contains 134 corporate garages of which 128 did not overcharge, i. e., we find that six (4.5%) of the garages overcharged their customers (see *Table 2.2*). This number is in accordance with Schneider (2012) who finds that in three out of 51 visits (or 6%) overcharging occurred.[25] Although 4.5%

[24]See, e. g., Johnson, D.L.: "6 tips to start your auto repair shop business today" (see `http://ezinearticles.com/?6-Tips-To-Start-Your-Auto-Repair-Shop-Business-Today&id=1176780`) or eHow: "How to open an auto repair shop" (see `http://www.ehow.com/how_2387498_open-auto-repair-shop.html`).

[25]The average amount overcharged was $32 per incident in the study by Schneider (2012). The sum of overchargings across all visits accounted for two percent of total charges.

Table 2.2: Descriptives.

Variable	Mean	Standard deviation	Min	Max	Obs.
Overcharging (= 1 if true)	0.045	0.208	0	1	134
Intense competition (= 1 if # of competitors is above median)	0.502	0.501	0	1	134
Critical financial situation (= 1 if true)	0.134	0.342	0	1	134
Competence (# of faults found out of 5)	4.239	1.125	0	5	134
Low reputation (= 1 if distance < 1500m)	0.224	0.418	0	1	134

overcharging cases might not seem to be a lot, the issue of overcharging turns out to be an important problem. The yearly turnover in the market for car repairs amounts to about 30 billion Euros in Germany alone (Zentralverband deutsches Kraftfahrtzeuggewerbe (Ed.), 2012). Following our data, the value of transactions where overcharging is involved would make up about 1.35 billion Euros per year which is far from negligible.

Table 2.2 also provides the descriptives for the four explanatory variables. 13.4% of the garages face a critical financial situation. About half of the garages face by construction of the variable an intense competition. The high competence (4.24 faults found out of 5) is due to the fact that the faults are all listed on the mechanics' checklists for inspections issued by all carmakers. Every fifth garage is close to the interstate and therefore faces low reputational concerns.

In order to provide a detailed characterization of the six garages that overcharged, *Table 2.3* lists the values for all four variables for each of these garages.[26] Note that there is considerable variation in the three variables critical financial situation,

[26]Note that the automobile club requested us not to publish names and addresses of the garages involved in the test. Therefore, garages are anonymous in *Table 2.3*.

Table 2.3: Characteristics of the garages that overcharge.

Garage	Intense competition	Critical financial situation	Competence	Low reputation
Garage 1	0	0	4	0
Garage 2	0	0	1	1
Garage 3	0	1	4	1
Garage 4	0	0	2	1
Garage 5	0	1	4	0
Garage 6	1	1	3	0

competence, and low reputation. The variable competition intensity, however, is almost separated. We will account for this quasi-separation in our data analysis by using a special type of regression analysis.

The correlations given in *Table 2.4* provide a first impression concerning the relationship between the different variables. All four explanatory variables prove to be correlated with the explained variable *overcharging*. Looking at the relationship between the explanatory variables, we observe that an intense competition is slightly correlated with low reputational concerns. Furthermore, a low competence is weakly correlated with a critical financial situation. This may be due to the fact that a garage with only a low competence attracts fewer customers than those garages with a high competence. Note, though, that the correlations between the variables amount to a maximum of 23.2% and are hence far from a collinear relationship.

Table 2.5 and *Figure 2.2* illustrate that the two groups—garages that do and do not overcharge—differ considerably in their characteristics: *Figure 2.2(a)* shows that garages that overcharge face an intense competition less often than those garages that do not overcharge. This difference in competition intensity is weakly significant (Mann Whitney U Test, two-tailed: $p = 0.096$). 50% of the garages that overcharge are in a critical financial situation whereas significantly fewer of those garages that do not overcharge have a critical financial background (11.7%, Mann Whitney U

Table 2.4: Correlations.

Variable	Over-charging	Intense competition	Critical fin. situation	Com-petence	Low reputation
Overcharging	1				
Intense competition	−0.144	1			
Critical fin. sit.	0.232	0	1		
Low reputation	0.104	0.232	−0.00507	1	
Competence	−0.239	−0.0266	−0.201	0.0875	1

Table 2.5: Mean comparisons between garages that did and did not overcharge.

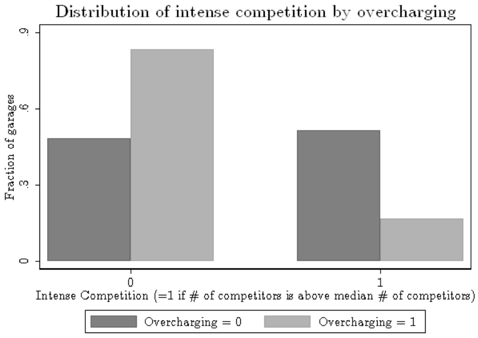
	Intense competition*	Critical financial situation***	Competence***	Low reputation
Overcharging = 1	0.167	0.500	3.000	0.500
Overcharging = 0	0.516	0.117	4.297	0.273

Mann Whitney U Test, two-tailed: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

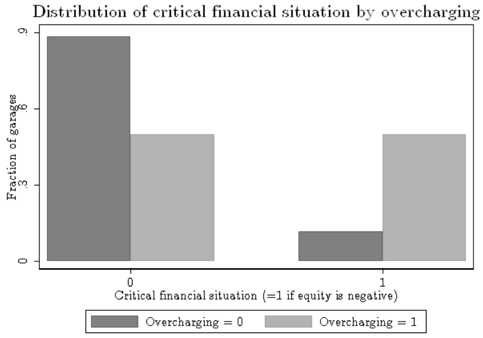
Test, two-tailed: $p = 0.007$; see also *Figure 2.2(b)*). The average competence of garages that overcharge is significantly lower than the average competence of those garages that do not overcharge (Mann Whitney U Test, two-tailed: $p = 0.003$; see also *Figure 2.2(c)*). *Figure 2.2(d)* suggests that garages that overcharge have low reputational concerns more often than garages that do not overcharge. However, this difference is not statistically significant (Mann Whitney U Test, two-tailed: $p = 0.231$).

2.4 Results

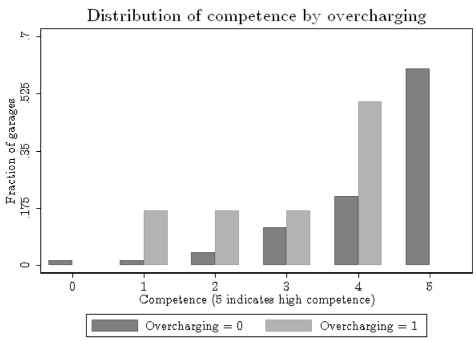
The small sample of our empirical analysis, the skewed distribution of our dependent variable, and the quasi-separation of the data with respect to competition intensity



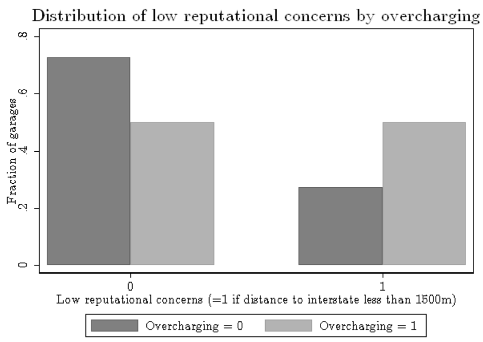
(a) Distribution of intense competition by overcharging.



(b) Distribution of critical financial situation by overcharging.



(c) Distribution of competence by overcharging.



(d) Distribution of low reputational concerns by overcharging.

Figure 2.2: Distribution of explanatory variables by overcharging.

represent a challenge concerning the deviation of meaningful conclusions. When addressing these issues, we make use of a well-established method—namely the Firth logit regression (Firth, 1993)—which is typically used in other research areas where small samples, skewed distribution of the dependent variable, and a quasi-separation are frequently observed phenomena. Most importantly, note that our results do *not* depend on the choice of the regression model used as we will show in the robustness checks (see section 2.5).

Let us shortly comment on the advantages of the Firth regression: the standard maximum likelihood estimation used in binary regression models assumes the sample to be large. As the sample size converges to infinity, the parameter estimates converge to the true parameter values. Hence, estimates may be biased in smaller samples. The Firth regression uses a penalized likelihood estimation removing the first-order bias that occurs due to the small sample (Heinze, 2006). The Firth approach also regularizes the data and thereby circumvents the separation problem (Zorn, 2005). Hence, the Firth regression always leads to finite parameter estimates which is not the case when using regressions based on the standard maximum likelihood estimation. The approach is frequently used in medical research[27] and has proven to outperform alternative small sample models such as the exact logistic regression (Heinze, 2006). Heinze (2006) highlights that for small samples "penalized likelihood confidence intervals for parameters show excellent behavior in terms of coverage probability and provide higher power than exact confidence intervals."

Given the four explanatory variables—competition intensity, financial situation, competence, and reputation—our Firth logit model is specified as follows:

$$\begin{aligned} firth\_logit(overcharging) = \beta_0 &+ \beta_1 intense\_competition \\ &+ \beta_2 critical\_financial\_situation \\ &+ \beta_3 competence + \beta_4 low\_reputation + \epsilon \end{aligned} \tag{2.4}$$

We report the results of the Firth regression in *Table 2.6.* We also present the results of the linear probability model in order to ease interpretation. To evaluate the model fit, we calculate McFadden's R^2 for the binary response models and the ordinary R^2 for the linear model. We choose to use McFadden's R^2 as a measure for the binary model fit as it can also be applied to the Firth logit regression. McFadden's R^2 is defined as $1 - L1/L0$ where $L1$ is the log-likelihood of the fully specified model and $L0$ is the log-likelihood of the null model. Interpreting $L0$ as the total sum of squares in linear regression analysis and $L1$ as the residual sum of squares, McFadden's R^2 provides a similar measurement for the model fit compared to the ordinary R^2 (Wooldridge, 2009). McFadden (1979) suggests that models with

[27] As an example, George et al. (2010) apply the Firth logit regression to the question of how a medication (phenylephrine) impacts spinal anesthesia-induced hypotension. Their work is based on a sample size of 45 test persons. Only nine test persons did not show a positive reaction to the medication.

Table 2.6: What drives fraud?

Overcharging	Firth Logit	OLS
Intense competition	-2.049^{**}	-0.078^{**}
(= 1 if # of competitors > median)	(1.040)	(0.035)
Critical financial situation	1.757^{**}	0.114^{**}
(= 1 if true)	(0.891)	(0.051)
Competence	-0.765^{**}	-0.041^{***}
(# of faults found out of 5)	(0.315)	(0.015)
Low reputational concerns	2.078^{**}	0.077^{**}
(= 1 if distance < 1500m)	(0.999)	(0.039)
Constant	-0.510	0.220^{***}
	(1.125)	(0.071)
McFadden R^2	0.412	–
R^2	–	0.142
Observations	134	134

Standard errors in parentheses, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
p-values are based on two-tailed tests.

an R^2 between 0.2 and 0.4 exhibit an excellent fit. The McFadden R^2 of our Firth regression amounts to 0.412 and is hence close to an excellent fit.

Let us next turn to the results.

Result 2.1. *Garages facing intense competition overcharge less often than those in a weakly competitive environment.*

In line with theory, we find that a high level of competition decreases the level of overcharging. According to the OLS estimates, a (highly) competitive environment decreases the probability of being overcharged by an expert by 7.8 percentage points. In fact, five out of the six garages that overcharge face a competition level that is lower than the median (see *Table 2.3*) whereas only every second garage that does not overcharge faces a competition level that is lower than the median (see *Table 2.5*).[28]

[28]Note that clearly, the effect of competition crucially depends on whether experts' and customers' interests with respect to fraudulent behavior are aligned or not (see footnote 7). In their empirical study, Bennett et al. (2013) find that competition among experts for vehicle emissions

Result 2.2. *A critical financial situation leads to a larger incentive to overcharge.*

Consistent with *Hypothesis 2.2*, we find that a critical financial situation increases a garage's incentive to overcharge. The OLS model estimates that a critical financial situation increases the probability of being overcharged by 11.4 percentage points. Garages in a critical financial situation overcharge more often compared to those with a solid financial background. In case overcharging is detected, the garage does not bear the costs of defrauding because it will file bankruptcy. On the other hand, if overcharging is not detected, the fraudulent behavior will help overcome the garages' financial difficulties.

Result 2.3. *A higher competence decreases the garages' incentive to overcharge.*

In line with *Hypothesis 2.3*, garages that exhibit high competence have a lower incentive to defraud their customers. The OLS regression results indicate that the probability of being overcharged decreases by 4.1 percentage points for each additional fault the garage detects.

Result 2.4. *Low reputational concerns increase the incentive to overcharge.*

Consistent with *Hypothesis 2.4*, the regression results show that low reputational concerns increase a garage's incentive to overcharge. The intuition is as follows: garages that have a low reputational concern, face many one-time interactions. Hence, they can overcharge their customers without hazarding a loss of future earnings. As recommended in Consumer Federation of America et al. (2011, p. 20), customers should "only do business with auto repair shops that you know and trust or that have good reputations based on other people's experiences. If you have any doubts about the diagnosis of your car's problem, bring it to another shop for a second opinion if possible." This statement is supported by our data. The OLS results

tests increases fraud. This is due to the fact that in their case, car owners whose cars are passed at higher rates due to fiercer competition may benefit from fraud as they save money on costly repairs. This, however, gives experts a greater incentive to generate a competitive advantage through illicit actions which raises the question whether competition is necessarily the ideal market structure in such an environment.

Table 2.7: Robustness against different models.

Overcharging	OLS	Logit	Probit	Scobit
Intense competition	-0.078^{**}	-2.593^{**}	-1.253^{**}	-2.539^{**}
(= 1 if # of competitors > median)	(0.035)	(1.262)	(0.605)	(1.162)
Critical financial situation	0.114^{**}	1.966^{*}	0.884^{*}	2.014^{**}
(= 1 if true)	(0.051)	(1.010)	(0.535)	(0.870)
Competence	-0.041^{***}	-0.887^{**}	-0.454^{**}	-0.835^{***}
(# of faults found out of 5)	(0.015)	(0.367)	(0.191)	(0.316)
Low reputational concerns	0.077^{**}	2.423^{**}	1.190^{**}	2.264^{**}
(= 1 if distance < 1500m)	(0.039)	(1.157)	(0.559)	(1.047)
Constant	0.220^{***}	−0.540	−0.282	−15.006
	(0.071)	(1.263)	(0.717)	(1878.318)
McFadden R^2	–	0.352	0.345	0.365
R^2	0.142	–	–	–
Observations	134	134	134	134

Standard errors in parentheses, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
p-values are based on two-tailed tests.

suggest that the probability of a garage overcharging its customer is increased by 7.7 percentage points if the garage has low reputational concerns.

2.5 Robustness Checks

Our results turn out to be extremely robust against alternative models such as the logit model with a regular maximum likelihood estimator, the probit, and the scobit regression (see *Table 2.7*). The latter accounts for the skewed distribution of the overcharging variable but is not significantly different from the logit regression. Significance levels of our explanatory variables remain practically unchanged when using these alternative models. The only decrease in a significance level from 5% to 10% occurs for the variable critical financial situation in the logit and probit model.

The results are also robust against choosing different parameters as cut-off points. In the above analysis, we measured the number of competitors within ten kilometers and then divided the garages in two categories: those facing less or more competitors than the median level. As *Table 2.8* shows, measuring the number of competitors within five or 20 kilometers instead of ten kilometers does not change our results. Note, however, that the significance level for the two variables competition intensity and critical financial situation decreases from 5% to 10% when varying the radius. Our results are also robust against including competition intensity as a continuous variable instead of using the dichotomized variable (see also *Table 2.8*). Yet, the significance level of competition intensity decreases from 5% to 10%. A possible explanation here might be that once a garage faces intense competition, the entry of further competitors does not influence the garage's overcharging decision.

Looking at the variable of low reputational concerns, *Table 2.8* shows that when considering those garages within 1000 or 2000 meters instead of 1500 meters to the next interstate as being close to the interstate, we do not obtain results any different from the above analysis. The significance level of competition intensity decreases from 5% to 10% when considering garages within 1000 meters. Similarly, the significance level of reputational concerns decreases from 5% to 10% when considering garages within 2000 meters to be close to the interstate.

Table 2.9 presents the results of our robustness checks with respect to alternative specifications. We control for yearly effects in order to ensure that the financial crisis does not affect garages' behavior. The results remain unchanged but the significance level of competition intensity and critical financial situation decreases from 5% to 10%. Furthermore, we show that whether a garage is an authorized or independent garage does not change any of our results. The significance level of a critical financial situation decreases from 5% to 10% when controlling for the years.

Table 2.8: Robustness against different cut-off points.

Overcharging	Firth logit competition 5k	Firth logit competition 20k	Firth logit competition continuous	Firth logit reputation 1000m	Firth logit reputation 2000m
Intense competition (= 1 if # of competitors within 5k > median)	-1.933^{*} (1.035)				
Intense competition (= 1 if # of competitors within 10k > median)				-1.759^{*} (1.006)	-2.327^{**} (1.075)
Intense competition (= 1 if # of competitors within 20k > median)		-1.844^{*} (1.019)			
Intense competition (continuous)			-0.014^{*} (0.008)		
Critical financial situation (= 1 if true)	1.546^{*} (0.861)	1.580^{*} (0.862)	1.876^{**} (0.901)	1.811^{**} (0.907)	1.864^{**} (0.887)
Competence (# of faults found out of 5)	-0.800^{**} (0.318)	-0.667^{**} (0.301)	-0.707^{**} (0.301)	-0.782^{**} (0.317)	-0.754^{**} (0.312)
Low reputational concerns (= 1 if distance < 1000m)				2.278^{**} (1.031)	
Low reputational concerns (= 1 if distance < 1500m)	1.985^{**} (0.991)	2.126^{**} (0.981)	2.274^{**} (1.026)		
Low reputational concerns (= 1 if distance < 2000m)					1.885^{*} (1.019)
Constant	-0.339 (1.136)	-0.891 (1.086)	-0.365 (1.163)	-0.563 (1.121)	-0.529 (1.120)
McFadden R^2	0.400	0.389	0.620	0.426	0.392
Observations	134	134	134	134	134

Standard errors in parentheses, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. p-values are based on two-tailed tests.

Table 2.9: Robustness against different specifications.

Overcharging	Firth logit	Firth logit controlling for authorized	Firth logit controlling for years
Intense competition	-2.049^{**}	-2.043^{**}	-1.956^{*}
(= 1 if # of competitors > median)	(1.040)	(1.036)	(1.160)
Critical financial situation	1.757^{**}	1.720^{*}	1.596^{*}
(= 1 if true)	(0.891)	(0.887)	(0.933)
Competence	-0.765^{**}	-0.747^{**}	-0.713^{**}
(# of faults found out of 5)	(0.315)	(0.312)	(0.317)
Low reputational concerns	2.078^{**}	2.017^{**}	2.286^{**}
(= 1 if distance < 1500m)	(0.999)	(0.984)	(1.056)
Authorized garage		1.037	
		(1.728)	
Year 2006			−0.260
			(1.555)
Year 2008			0.179
			(1.295)
Year 2009			−1.190
			(1.397)
Constant	−0.510	−0.507	−0.257
	(1.125)	(1.119)	(1.226)
McFadden R^2	0.412	0.375	0.426
Observations	134	134	134

Standard errors in parentheses, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
p-values are based on two-tailed tests.

Considering the full dataset of 297 corporate and non-corporate garages that do not belong to a chain, the Firth regression shows that a low competence increases the garages' incentive to overcharge. Low reputational concerns and an intense competition do not have a significant influence on the incentive to overcharge in the full dataset. This may be due to three reasons: firstly, we have shown that for corporate garages, the financial situation has a significant impact on the garages' overcharging behavior. As we are not able to control for the financial situation in

Table 2.10: Robustness against full sample

Overcharging	Firth logit
Intense competition	-0.391
(= 1 if # of competitors > median)	(0.627)
Competence	-0.415^{**}
(# of faults found out of 5)	(0.200)
Low reputational concerns	-0.011
(= 1 if distance < 1500m)	(0.664)
Constant	-1.530^{*}
	(0.789)
McFadden R^2	0.106
Observations	297

Standard errors in parentheses, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
p-values are based on two-tailed tests.

the unrestricted dataset, the estimates suffer from omitted variable bias. It may be this bias leading to lower parameter estimates and thus to a lower significance level. Secondly, non-corporate garages are more heterogeneous making it more difficult to identify effects. The McFadden R^2 amounts to 0.412 for the Firth regression on corporate garages whereas the indicator for the model fit drops to 0.106 when considering the whole sample. A third reason why only competence remains to have a significant influence on the level of overcharging when considering the whole sample may be that non-corporate garages react less sensitive to different levels of competition and to the number of one-time interactions than corporate garages. A possible explanation here is that non-corporate garages are often owned by self-employed mechanics. Many of those mechanics have worked for decades in their garage and do not adjust their overcharging behavior anymore when the competitive environment changes or new interstates are built close to their garage.

2.6 Conclusion

Making use of a field study, we analyze the impact of car repair shops' reputational concerns, their financial situation, the degree of market competition, and the garages' competence on their incentive to overcharge. We focus on corporate garages in our analysis and find in—accordance with theory—that firms that care little about their reputation and those that struggle with a critical financial situation have a greater incentive to defraud their customers. On the other hand, firms with a high competence are less likely to overcharge. While Dulleck et al. (2011) do not find support for an effect of competition on the probability of overcharging in their experimental study, we show that in a more competitive environment, the expert's incentive to overcharge decreases. As such, our results provide field evidence for many of the aspects often found in recommendations by consumer-protection agencies.

There are two limitations of our study: firstly, the study focuses on 134 corporate garages of which only six overcharged. The low number of observations and the skewed distribution of overcharging challenge identification. Using Firth regression, we account for these challenges and find considerably robust results that do not depend on the choice of method but hold across all conventional regression methods. The second limitation is that when considering the full dataset of 297 garages, we find that only competence remains to have a significant influence on the level of overcharging. Competition intensity and reputational concerns no longer have a significant influence on the garages' incentive to overcharge.

On a general perspective, our results may provide insights into and testable hypotheses for the functioning of other credence goods markets. For example, applying our results to the health care market[29], a high physician density should reduce the physicians' incentive to overcharge. Additionally, general practitioners with repeated

[29] The health care market is the largest credence goods market in most industrialized economies, making up about 10% of the GDP (OECD, 2011). For the US, the FBI estimates that up to 10% of those expenditures are due to fraudulent behavior (Federal Bureau of Investigation, 2007).

patient interaction should face a lower incentive to overcharge than specialists who are often only consulted once. Furthermore, our results may also provide important implications for the comparison across different credence goods markets. Whereas the cab market is characterized by one-time interactions, the market for legal advice is usually characterized by repeated interaction. In light of our analysis, we should expect more overcharging for taxi rides than for legal advice. Whether this is indeed the case is left for analysis in future studies.

2.7 Appendix A: A Model of the Car Repair Market: Theoretical Predictions

In the market with homogeneous customers and experts described in section 2.2, the following result is obtained:[30]

Proposition 2.1. *There exists a symmetric weak perfect Bayesian equilibrium with the following characteristics:*

(i) *experts set prices* $p_L = c_L + \Delta$ *and* $p_H = c_H > c_L + \Delta$ *(where* $\Delta > 0$ *is a markup);*

(ii) *experts always recommend the major treatment if the customer has the major problem and they recommend the major treatment with probability* $x \in (0, 1)$ *if the customer has the minor problem (overcharging);*

(iii) *customers at their first visit always accept a minor recommendation and accept a major recommendation with probability* $y \in (0, 1)$ *and customers who visit a second (different) expert accept both recommendations with certainty; and*

(iv) *a customer who accepts a recommendation always gets sufficient treatment.*

Proof. Note that result (iv) is straightforward: due to liability, experts cannot undertreat their customers. Moreover, from the prices given in the proposition it follows that the cost differential satisfies $c_H - c_L > \Delta$, i.e., experts have no incentive to overtreat their customers.

In order to fully characterize an equilibrium with the characteristics mentioned in the proposition, consider the expert's recommendation decision given the customer's acceptance decision as specified in the proposition. As mentioned in the main text, in equilibrium, an expert consulted by a customer with a minor problem must be

[30] The market and the insights presented here represent one of the cases discussed by Dulleck and Kerschbamer (2006) (see part (i) of their *Lemma 6* and the respective proof). The arguments to derive the first result closely follow their analysis.

indifferent between recommending the minor and major treatment, i.e.,

$$p_L - c_L = \frac{y + x(1-y)}{1 + x(1-y)} \left(p_H - c_L\right). \tag{2.5}$$

Hence, the expert makes a strictly positive profit with the minor recommendation with certainty. The payoff from recommending the major treatment equals a lottery: if the recommendation is accepted which happens with a probability smaller than one, the experts makes a profit that is higher than for the minor recommendation; however, if the recommendation is not accepted, the payoff is equal to zero.

Next, consider customers' acceptance decision: again as highlighted in the main text, a customer given the major recommendation must be indifferent between rejecting and accepting the diagnosis, i.e.,

$$d = \frac{x(1-h)}{h + x(1-h)}(1-x)(p_H - p_L). \tag{2.6}$$

Hence, the additional costs of searching for a second opinion d must equal expected savings from visiting a second expert firm (right-hand side). With probability $x(1-h)/(h + x(1-h))$, the customer has a minor problem given a major recommendation by the first expert. With probability $1-x$, the second expert is honest and recommends the minor treatment which means that the customer saves the cost differential $p_H - p_L$ compared to the first recommendation. Note that here, it becomes clear why a third visit does not pay off for a customer who is indifferent between accepting and rejecting a high-treatment recommendation on her first visit: if she receives a high-treatment recommendation from a second expert, the probability that she actually only needs the minor treatment is lower compared to the first visit.

Furthermore, a customer who gets the minor recommendation always accepts. This means that experts always recommend the major treatment if the the customer the the major problem as $p_L < c_H$.

Hence, for exogenously fixed prices $p_L = c_L + \Delta$ and $p_H = c_H > c_L + \Delta$ as well as for a markup Δ such that both the recommendation probability x and the acceptance probability y satisfy the compatibility constraints given by equations (2.5) and (2.6) and lie in between zero and one, the situation described in parts (i)–(iv) in the proposition is indeed part of a perfect Bayesian equilibrium.

Now consider the case where experts set prices within the given range. Denote by $\bar{x}$ ($\underline{x}$) the probability that an expert recommends the major treatment when the customer has the minor (major) problem. Furthermore, a customer who is recommended the major (minor) treatment believes that she has the major problem with probability $\bar{\mu}$ ($\underline{\mu}$). Accordingly, $\bar{y}$ ($\underline{y}$) denotes the probability that a customer accepts the recommendation of a major (minor) treatment. Last, a customer incurs expected costs of $k = d + (1-h)(1-x)(c_L + \Delta) + (h + (1-h)x)c_H > 0$ when she follows the proposed equilibrium strategy and an experts make a profit of $\pi = (1-h)(1+x(1-y))\Delta > 0$ per customer when they stick to the proposed equilibrium strategy.

As far as customers' beliefs are concerned, suppose that beliefs are correct whenever expert charge those prices given in the proposition, i.e., $\bar{\mu}(p_L, p_H) = (h + x^2(1-h))/(h + x(1-h))$ and $\underline{\mu}(p_L, p_H) = x(1-h)/(h + x(1-h))$. Moreover, suppose that for out-of-equilibrium beliefs, it holds that (i) $\bar{\mu}(p_L, p_H) = 1$ and $\underline{\mu}(p_L, p_H) = 0$ if and only if $p_L \leq d + (1-x)(c_L + \Delta) + xc_H$ and $p_H \in [c_H, c_H + d)$ and (ii) $\bar{\mu}(p_L, p_H) = h$ and $\underline{\mu}(p_L, p_H) = 0$ otherwise.

Next, consider the following acceptance decisions: (i) $\underline{y}(p_L, p_H) = 1$ if and only if $p_L \leq d + (1-x)(c_L + \Delta) + xc_H$ and $\underline{y}(p_L, p_H) = 0$ otherwise and (ii) $\bar{y}(p_L, p_H) = 1$ if and only if either $p_L \leq d + (1-x)(c_L + \Delta) + xc_H$ and $p_H \leq c_H + d$ or $p_L > d + (1-x)(c_L + \Delta) + xc_H$ and $p_H \leq k$ and $\bar{y}(p_L, p_H)$ else.

Suppose further that a deviating expert always recommends the major treatment (i.e., $\underline{x}(p_L, p_H) = \bar{x}(p_L, p_H) = 1$), a customer never consults a deviating expert, and the experts' price-posting strategy stipulates that they never deviate to set prices different from the ones given in the proposition.

To check whether the equilibrium candidate characterized above is a weak perfect Bayesian equilibrium, consider first the acceptance decisions: if a single expert deviates, the proposed price vector is still available because there is at least one remaining expert offering treatment services at these prices. Compared with expected cost k, a customer who believes that she has the minor (major) problem with certainty faces lower (higher) costs equal to $d + (1 - x)(c_L + \Delta) + xc_H$ $(d+c_H)$. Hence, customers' acceptance decisions are optimal. Given these decisions, $\underline{x}(p_L, p_H) = \bar{x}(p_L, p_H) = 1$ is optimal for a deviating expert as either $\bar{y}(p_L, p_H) = 1$ and $p_H \geq c_H$ or $\underline{y}(p_L, p_H) = \bar{y}(p_L, p_H) = 0$. In light of this recommendation policy and the observation that $p_H \geq c_H$, customers indeed rather stay away from deviating experts, i.e., their deviating profit is zero. □

Impact of the number of firms

We consider the following adaptation of the initial market setting to analyze how a change in the number of experts influences the incentives to overcharge: suppose that an increase in the number of firms n leads to a decrease in search costs $d(n)$ as customers have to spend less time and effort searching for suitable experts.[31]. In this case, we can readily state the following lemma:

Lemma 2.1. *All else equal, an increase in the number of expert firms active in the market reduces their incentive to overcharge.*

Proof. In this case, the initial indifference condition regarding a customer's acceptance decision given in (2.6) changes to

$$d(n) + \frac{x(1-x)(1-h)}{h + x(1-h)} p_L + \left(1 - \frac{x(1-x)(1-h)}{h + x(1-h)}\right) p_H = p_H. \tag{2.7}$$

Note that the left-hand side of equation (2.7) is lower than the one in equation (2.6). This means that customers find a second expert more easily and hence, the acceptance probability y of a major recommendation goes down. This in turn leads

[31]For example, if experts are horizontally differentiated, customers have to incur less transportation costs to reach a second expert when the number of experts in the market goes up.

to a decrease in the probability that an expert firm dishonestly recommending the major treatment actually gets the business. More precisely, let $\chi := (y + x(1 - y))/(1 + x(1 - y))$. Then, $\partial\chi/\partial y = 1/(1 + x(1 - y))^2 > 0$. As a consequence, the scope for fraud is reduced as n increases because cheating becomes less profitable. □

Impact of the financial situation

In order to analyze the effect of an expert firm's financial situation on the incentives to overcharge, consider the following change to the situation described above: different from the initial setting, suppose that firms have identical fixed costs f to run their business but are heterogeneous regarding their financial assets. There are two groups of firm: firms in the first group need to attract customers as they only have limited resources left to pay their fixed costs f. Importantly, these firms only pay the fixed cost if they attract a customer. If they do not, they go bankrupt and receive a payoff of zero due to their limited liability. Firms in the second group have a much sounder financial background which means that they survive the current period even if they incur fixed costs without serving any customer. The following lemma takes a closer look at firms' incentives to defraud their customers in both groups:

Lemma 2.2. *All else equal, an expert firm which is in a critical financial situation is more likely to overcharge for its services.*

Proof. In this case, the initial incentive-compatibility constraint by equation (2.5) changes for an expert firm that is in financial distress to

$$p_L - c_L - f = \frac{y + x(1-y)}{1 + x(1-y)} \left(p_H - c_L - f\right). \tag{2.8}$$

Analogously, the incentive-compatibility constraint for the firm with the strong financial background must be equal to

$$p_L - c_L - f = \frac{y + x(1-y)}{1 + x(1-y)} \left(p_H - c_L - f\right) - \left(1 - \frac{y + x(1-y)}{1 + x(1-y)}\right) f. \tag{2.9}$$

Plugging constraint (2.8) into constraint (2.9) gives

$$\frac{y+x(1-y)}{1+x(1-y)}(p_H-c_L-f) > \frac{y+x(1-y)}{1+x(1-y)}(p_H-c_L-f) - \frac{1-(y+x(1-y)}{1+x(1-y))}f.$$

This means that whenever the incentive-compatibility constraint is satisfied for the financially weak expert firm, it is also satisfied for the financially strong firm. As a result, the latter has a lower incentive to defraud its customers as it finds it more profitable to recommend the minor treatment whenever it is needed. □

Impact of the expert's competence

Last, we analyze the effect of an expert firm's competence on the incentives to overcharge. To this end, consider the following change to the above framework. Again, there are two groups of firms. Firms in the two groups are heterogeneous with respect to their competence. The firms in the first group are of low competence and firms still incur costs c_L and c_H for the low and the major treatment. On the other hand, the firms of high competence in the second group can offer these services at lower costs of $c_L - \gamma$ and $c_H - \gamma$, respectively. Given this setup, we can state the following lemma:

Lemma 2.3. *All else equal, a high-competence firm is less likely to overcharge compared to its low-competence competitor.*

Proof. Note first that the incentive-compatibility constraint for the low-competence expert firm is the same as in the original setting and given by expression (2.5). The incentive-compatibility constraint for the high-competence firm equals

$$p_L - (c_L - \gamma) = \frac{y+x(1-y)}{1+x(1-y)}\left(p_H - (c_L - \gamma)\right). \tag{2.10}$$

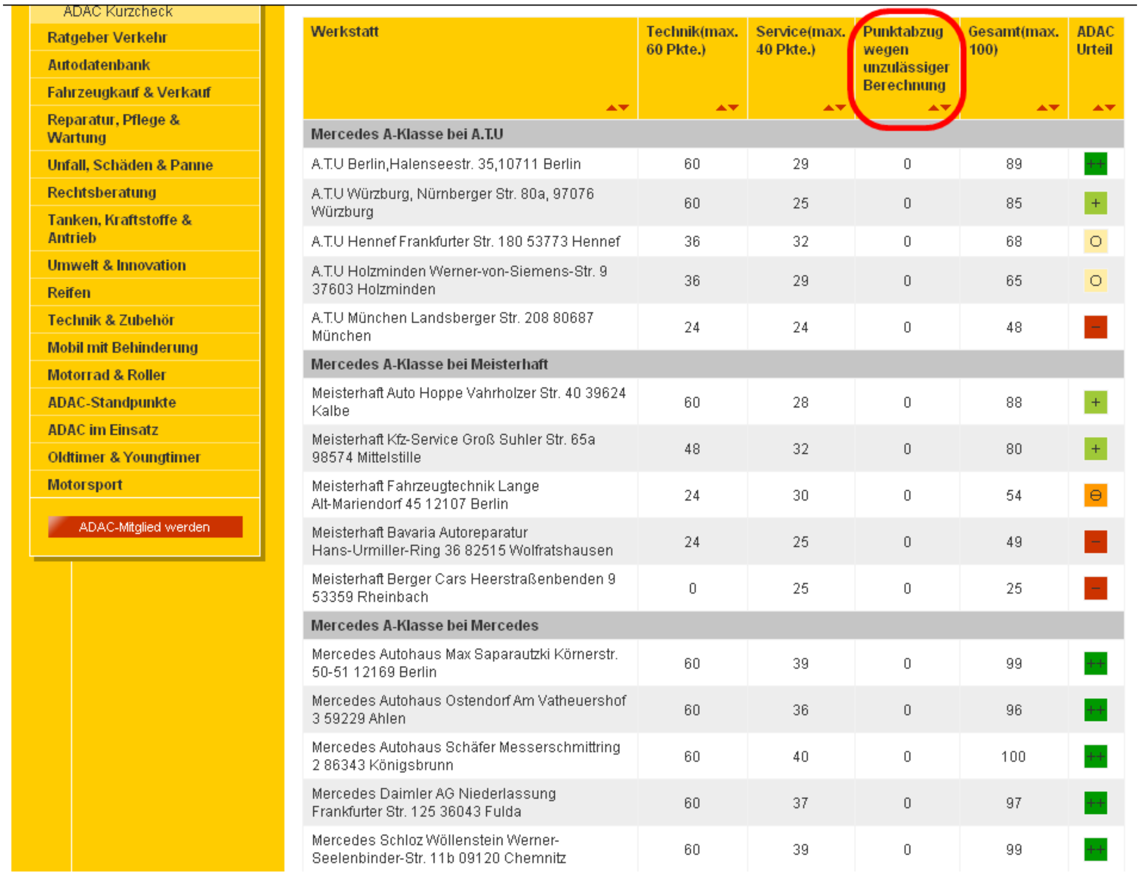
Plugging constraint (2.5) into constraint (2.10) gives

$$\frac{y+x(1-y)}{1+x(1-y)}(p_H-c_L)+\gamma > \frac{y+x(1-y)}{1+x(1-y)}(p_H-c_L+\gamma).$$

We can thus conclude that the high-competence firm has a lower incentive to defraud its customers. $\square$

2.8 Appendix B: Screenshots of Data Collection

2.8.1 Overcharging



ADAC Kurzcheck
Ratgeber Verkehr
Autodatenbank
Fahrzeugkauf & Verkauf
Reparatur, Pflege & Wartung
Unfall, Schäden & Panne
Rechtsberatung
Tanken, Kraftstoffe & Antrieb
Umwelt & Innovation
Reifen
Technik & Zubehör
Mobil mit Behinderung
Motorrad & Roller
ADAC-Standpunkte
ADAC im Einsatz
Oldtimer & Youngtimer
Motorsport
ADAC-Mitglied werden

Werkstatt	Technik(max. 60 Pkte.)	Service(max. 40 Pkte.)	Punktabzug wegen unzulässiger Berechnung	Gesamt(max. 100)	ADAC Urteil
Mercedes A-Klasse bei A.T.U					
A.T.U Berlin,Halenseestr. 35,10711 Berlin	60	29	0	89	++
A.T.U Würzburg, Nürnberger Str. 80a, 97076 Würzburg	60	25	0	85	+
A.T.U Hennef Frankfurter Str. 180 53773 Hennef	36	32	0	68	O
A.T.U Holzminden Werner-von-Siemens-Str. 9 37603 Holzminden	36	29	0	65	O
A.T.U München Landsberger Str. 208 80687 München	24	24	0	48	–
Mercedes A-Klasse bei Meisterhaft					
Meisterhaft Auto Hoppe Vahrholzer Str. 40 39624 Kalbe	60	28	0	88	+
Meisterhaft Kfz-Service Groß Suhler Str. 65a 98574 Mittelstille	48	32	0	80	+
Meisterhaft Fahrzeugtechnik Lange Alt-Mariendorf 45 12107 Berlin	24	30	0	54	⊖
Meisterhaft Bavaria Autoreparatur Hans-Urmiller-Ring 36 82515 Wolfratshausen	24	25	0	49	–
Meisterhaft Berger Cars Heerstraßenbenden 9 53359 Rheinbach	0	25	0	25	–
Mercedes A-Klasse bei Mercedes					
Mercedes Autohaus Max Saparautzki Körnerstr. 50-51 12169 Berlin	60	39	0	99	++
Mercedes Autohaus Ostendorf Am Vatheuershof 3 59229 Ahlen	60	36	0	96	++
Mercedes Autohaus Schäfer Messerschmittring 2 86343 Königsbrunn	60	40	0	100	++
Mercedes Daimler AG Niederlassung Frankfurter Str. 125 36043 Fulda	60	37	0	97	++
Mercedes Schloz Wöllenstein Werner-Seelenbinder-Str. 11b 09120 Chemnitz	60	39	0	99	++

Figure 2.3: Data collection on the overcharging measurement. Source: `http://www.adac.de`, accessed on January 17, 2012.

2.8.2 Intense Competition

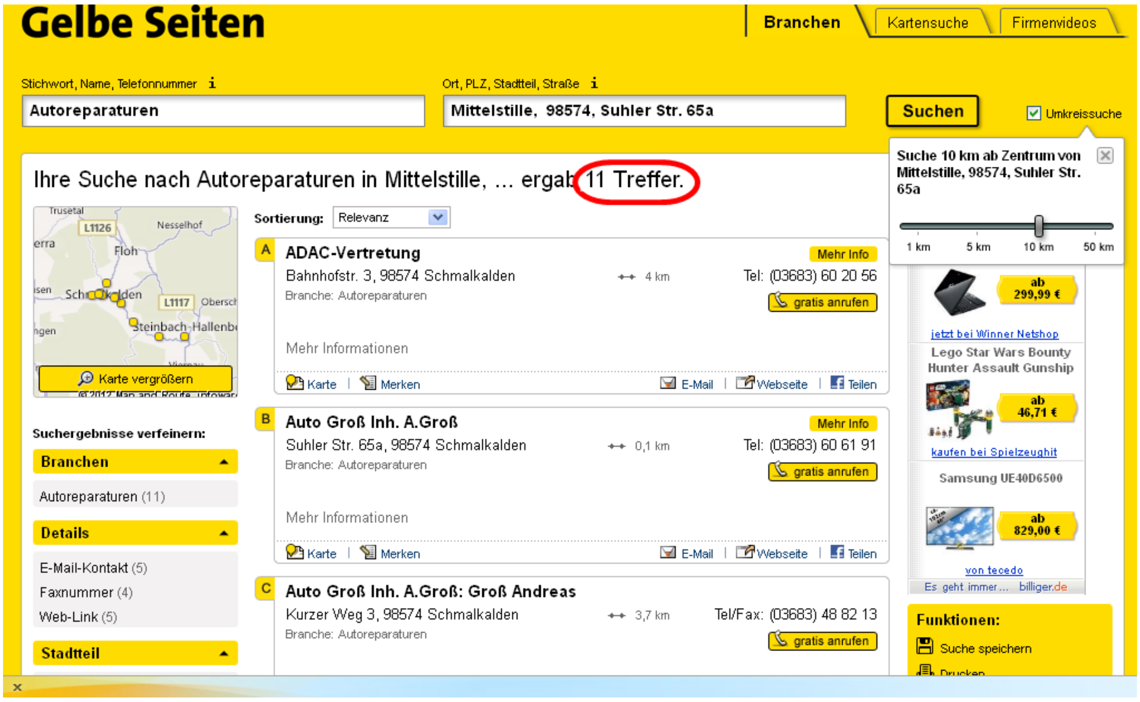


Figure 2.4: Data collection on the competition measurement. Source: `http://www.gelbeseiten.de`, accessed on January 17, 2012.

2.8.3 Financial Situation

Kfz-Service Andreas Groß GmbH

Mittelstille

Jahresabschluss zum 31.12.2006

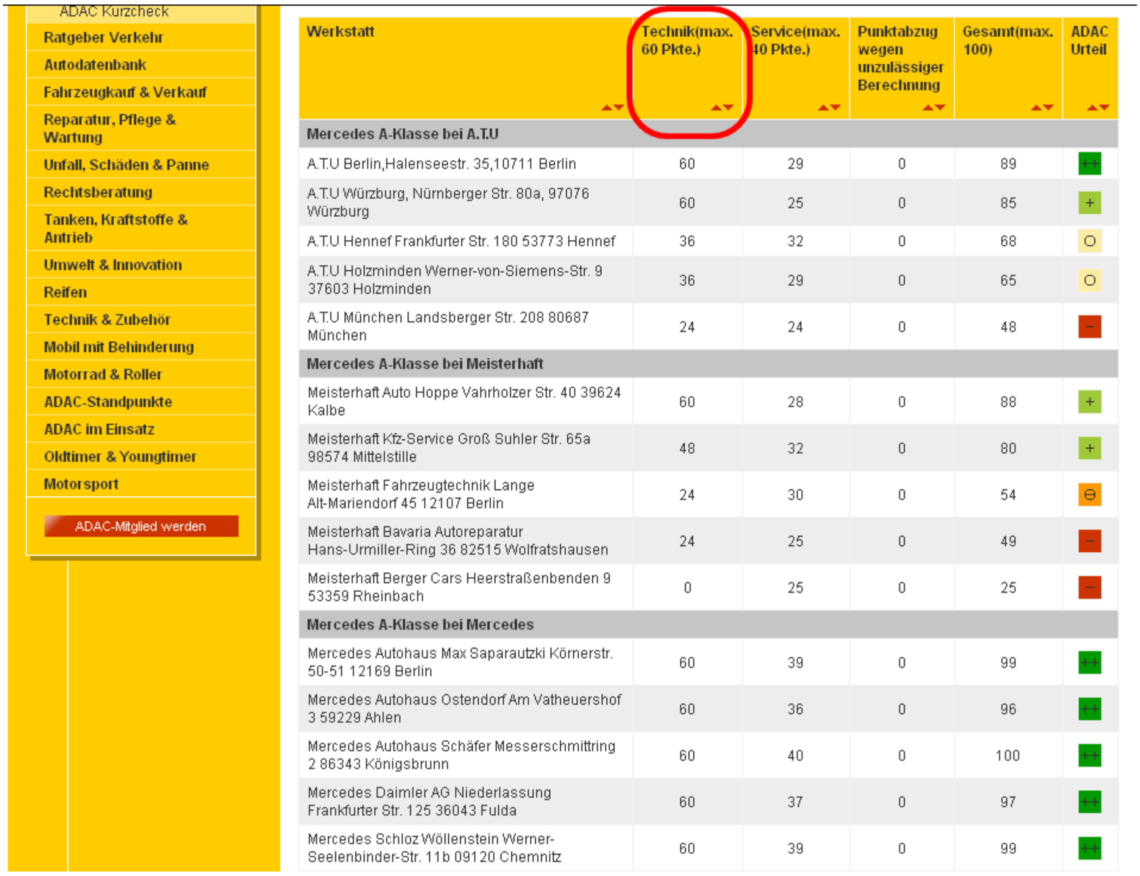
Bilanz

Aktiva	31.12.2006 EUR
A. Anlagevermögen	16.220,50
I. Sachanlagen	16.220,50
B. Umlaufvermögen	264.270,91
I. Vorräte	114.367,52
II. Forderungen und sonstige Vermögensgegenstände	138.962,74
III. Kassenbestand, Bundesbankguthaben, Guthaben bei Kreditinstituten und Schecks	10.940,65
Bilanzsumme, Summe Aktiva	280.491,41

Passiva	31.12.2006 EUR
A. Eigenkapital	39.142,05
I. gezeichnetes Kapital	26.000,00
II. Gewinnvortrag	10.919,03
III. Jahresüberschuss	2.223,02
B. Sonderposten mit Rücklageanteil	4.000,00
C. Rückstellungen	8.415,47
D. Verbindlichkeiten	228.933,89
Bilanzsumme, Summe Passiva	280.491,41

Figure 2.5: Data collection on the financial situation. Source: `http://www.bundesanzeiger.de`, accessed on January 17, 2012.

2.8.4 Competence



ADAC Kurzcheck
Ratgeber Verkehr
Autodatenbank
Fahrzeugkauf & Verkauf
Reparatur, Pflege & Wartung
Unfall, Schäden & Panne
Rechtsberatung
Tanken, Kraftstoffe & Antrieb
Umwelt & Innovation
Reifen
Technik & Zubehör
Mobil mit Behinderung
Motorrad & Roller
ADAC-Standpunkte
ADAC im Einsatz
Oldtimer & Youngtimer
Motorsport
ADAC-Mitglied werden

Werkstatt	Technik(max. 60 Pkte.)	Service(max. 40 Pkte.)	Punktabzug wegen unzulässiger Berechnung	Gesamt(max. 100)	ADAC Urteil
Mercedes A-Klasse bei A.T.U					
A.T.U Berlin,Halenseestr. 35,10711 Berlin	60	29	0	89	++
A.T.U Würzburg, Nürnberger Str. 80a, 97076 Würzburg	60	25	0	85	+
A.T.U Hennef Frankfurter Str. 180 53773 Hennef	36	32	0	68	O
A.T.U Holzminden Werner-von-Siemens-Str. 9 37603 Holzminden	36	29	0	65	O
A.T.U München Landsberger Str. 208 80687 München	24	24	0	48	−
Mercedes A-Klasse bei Meisterhaft					
Meisterhaft Auto Hoppe Vahrholzer Str. 40 39624 Kalbe	60	28	0	88	+
Meisterhaft Kfz-Service Groß Suhler Str. 65a 98574 Mittelstille	48	32	0	80	+
Meisterhaft Fahrzeugtechnik Lange Alt-Mariendorf 45 12107 Berlin	24	30	0	54	⊖
Meisterhaft Bavaria Autoreparatur Hans-Urmiller-Ring 36 82515 Wolfratshausen	24	25	0	49	−
Meisterhaft Berger Cars Heerstraßenbenden 9 53359 Rheinbach	0	25	0	25	−
Mercedes A-Klasse bei Mercedes					
Mercedes Autohaus Max Saparautzki Körnerstr. 50-51 12169 Berlin	60	39	0	99	++
Mercedes Autohaus Ostendorf Am Vatheuershof 3 59229 Ahlen	60	36	0	96	++
Mercedes Autohaus Schäfer Messerschmittring 2 86343 Königsbrunn	60	40	0	100	++
Mercedes Daimler AG Niederlassung Frankfurter Str. 125 36043 Fulda	60	37	0	97	++
Mercedes Schloz Wöllenstein Werner-Seelenbinder-Str. 11b 09120 Chemnitz	60	39	0	99	++

Figure 2.6: Data collection on the competence measure. Source: `http://www.adac.de`, accessed on January 17, 2012.

2.8.5 Low Reputation

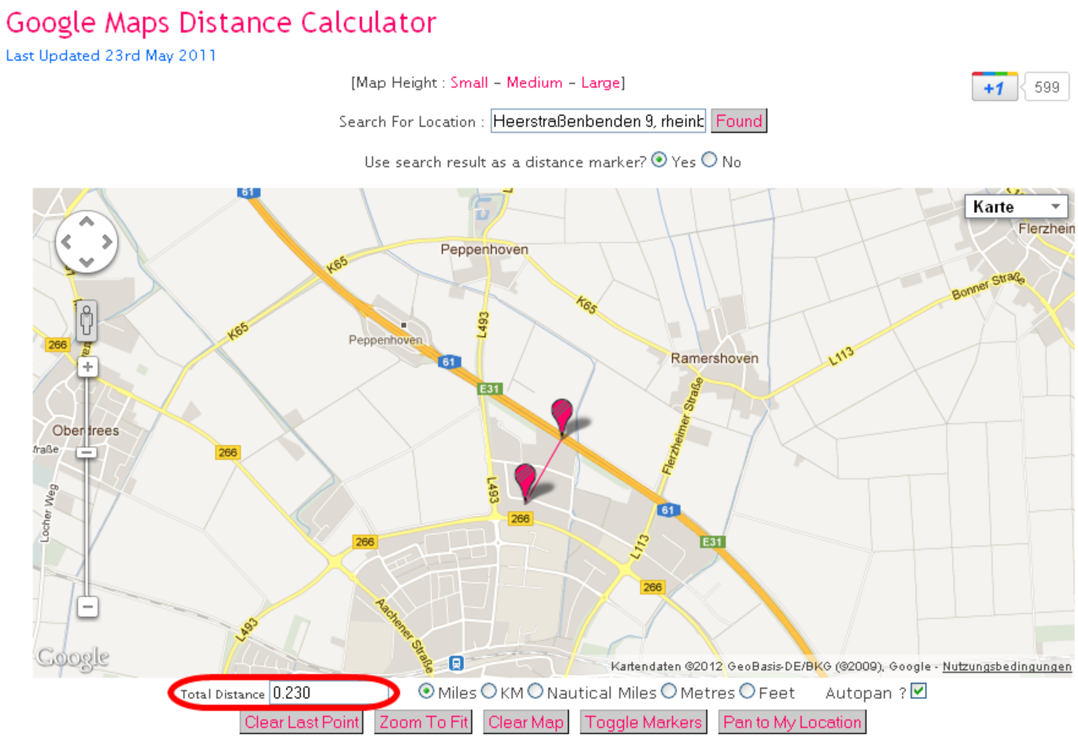


Figure 2.7: Data collection on the reputation measure. Source: `http://www.daftlogic.com`, accessed on January 17, 2012.

3 Fraud and Other-Regarding Preferences in a Market for Credence Goods

We introduce other-regarding preferences in a credence goods model applied to the health care market. Physicians act as experts and do not only care about their monetary payoff but also about treating their patients honestly. Patients, on the other hand, face trust costs when they anticipate to be defrauded. We study the impact of those other-regarding preferences on the level of fraud in the market. Contrary to intuition, we show that the level of fraud does not necessarily decrease in equilibrium when introducing these other-regarding preferences but may even rise. Welfare increases with experts' honesty but reacts ambiguously to increased patients' trust costs.

3.1 Introduction

Medical services are a prime example of credence goods (Darby and Karni, 1973). The patient notices that she is ill but she does not know which disease she suffers from. The physician performs a diagnosis and can thereby identify whether the patient suffers from a minor or a major disease. If the treatment is not verifiable to the patient, the physician can exploit his informational advantage by providing the minor treatment for the minor disease and charging for performing the major treatment, i.e., the physician can overcharge his patient. This paper analyzes the impact of other-regarding preferences on the level of overcharging in a credence goods market.

Previous work on credence goods has focused on the physicians' and patients' monetary incentives (see, e. g., Dulleck and Kerschbamer, 2006; Emons, 1997; Wolinsky, 1993). However, physicians do not only care about monetary incentives but also about honesty. Dulleck et al. (2011) show in a large scale experiment that experts like physicians base their decisions to a "considerable extent" on norms like honesty. Even in a situation where theory would predict that experts always defraud their customers, only half of them actually do so. In a follow-up study, Beck et al. (2011) find that experts feel guilty when defrauding their customers. In addition to the honesty concerns experts face, physicians take the Oath of Hippocrates swearing that they will only act for the benefit of the sick (Pieper and Thurnherr, 1998). The oath might give the physicians a reason to behave even more honestly than other experts do.[32] Patients, on the other hand, do not only care about their monetary costs of the treatment. Tibandebage and Mackintosh (2005) show that in Tanzania, where medical expenses are completely paid out of pocket, patients lose trust in the honesty of their physician when they expect to be defrauded. Fehr and Gächter (2002) even show that patients are willing to pay in order to punish the physician when they presume to be cheated on.

This work introduces other-regarding preferences in a credence goods model. Physicians face *conscience costs* in our model when they overcharge their patients. Patients face *trust costs* when being charged a major treatment because they anticipate that they may have been defrauded. We study the impact of these other-regarding preferences on the level of overcharging when patients can consult a second physician. A first intuition would lead to the conclusion that introducing such preferences decreases the level of overcharging in the market. We show that this intuition does not necessarily hold in equilibrium. An increase in *trust costs* can

[32] Ahlert et al. (2008) and Hennig-Schmidt and Wiesen (2010) provide experimental evidence for a more honest behavior of prospective physicians compared to other subject pools.

lead to more searches for a second opinion in equilibrium, making it more attractive for physicians to defraud their patients.

The problem of credence goods is not limited to the health care market: the cab ride in an unknown city (Balafoutas et al., 2013), the purchase of a highly technical product, and a lawyer's advice are other examples. The largest credence goods market in most economies is the health care market though. In the US, about 2.4 trillion USD are spent per year for health care (OECD, 2011), up to 10% of which are estimated to be due to fraud (Federal Bureau of Investigation, 2007). We therefore focus on the health care market in our paper.

Related to our work is Liu (2011). She shows that two types of equilibria may exist in a credence goods market with a possibly selfish or conscientious monopolistic expert: a uniform price equilibrium in which the selfish expert free rides on the conscientious and a non-uniform price equilibrium in which the customer can identify the type of expert by the price list. In contrast to Liu (2011), we consider homogeneous physicians with other-regarding preferences. We focus on the influence of these preferences on the level of overcharging in the market rather than on the separation of physician types.[33]

The remainder of the article is as follows: In section 3.2 we introduce the model and other-regarding preferences. In section 3.3 we perform an equilibrium analysis. The last section concludes and highlights implications.

3.2 Model

The model is built upon Wolinsky (1993): There is a continuum of patients and a large number of physicians in the health care market. Each patient either suffers

[33]In our model, patients can search for a second opinion in order to discipline physicians. Contrary to Liu (2011), we assume prices to be exogenously given as most medical markets are highly regulated. Furthermore, we do not allow for a rejection of interaction by either party. This is because in most countries physicians are legally obliged to treat their patients. Patients, often, do not have a choice of whether to get treated or not because diseases aggravate. While Liu (2011)'s results are driven by this rejection of an interaction, our results are driven by the patients' search for a second opinion.

from a minor or a major disease. The patient notices that she is ill but does not know from which disease she suffers. It is common knowledge that the patient has a major disease with probability h or a minor disease with probability $1-h$. Physicians are able to costlessly diagnose the disease. If the patient accepts the diagnosis, the physician treats the patient. Treating the major disease induces costs of c_H for the physicians, treating the minor disease of $c_L < c_H$. Patients derive a utility of B when they are treated.

Patients cannot verify the treatment. Therefore, physicians can overcharge their patients given that they suffer from a minor disease. Undertreatment is ruled out because physicians are assumed to be liable. In order to discipline physicians, patients can search for a second opinion. Searching for a physician induces costs d for the patient. The physician does not know whether a patient is on her first or second visit. It is assumed that B is large enough so that getting treated makes the patient always better off than no treatment.

In contrast to Wolinsky (1993) physicians and patients do not only follow monetary incentives but also care about honesty and trust. Ariely and Mazar (2006) show that experts like physicians trade-off between pursuing their own financial goals and being honest. They find that people "possess internal reward mechanisms for honesty" that "limit dishonest behavior." Physicians therefore face *conscience costs* in our model when they overcharge their patient. Gneezy (2005) shows that the non-monetary costs of defrauding increase with an increasing loss of the other party. Consequently, we model a physician's *conscience costs* by subtracting a share β of his fraudulent profit.

Patients, on the other hand, lose trust in their physician when they are confronted with a major diagnosis. This is because they anticipate that they may have been overcharged (Tibandebage and Mackintosh, 2005). We therefore add *trust costs* of α weighted by the probability of being defrauded. Thus, an increase in the market level of fraud implies higher *trust costs*.

We assume prices for the treatments to be exogenously given, e. g. by a regulator. The price for the major treatment amounts to $p_H = c_H$, the price for the minor treatment to $p_L = c_L + e$. e represents the mark-up on the low cost treatment and ensures an incentive for the physician to treat his patient honestly. The reason for assuming fixed prices is that in many medical markets we can indeed observe exogenously given prices, either because of a bargaining process on a central level (as in the US) or by legal regulations (as e. g. in Germany) (Sülzle and Wambach, 2005).

The patient is not insured against health risks in our model. Consequently, the treatment costs are completely borne by the patient herself. The impact of insurance is discussed in section 3.2.1. Health insurance does not change any of our results as long as the patient faces a co-payment.

The game structure is as follows:

1. Nature determines the severity of the patient's disease: With probability h the patient suffers from a major disease, with probability $1 - h$ from a minor disease.

2. The patient chooses a physician and incurs search costs d.

3. The physician learns the patient's disease. Given that the patient suffers from a minor disease, the physician diagnoses a minor disease with probability $1-x$ ($x \in [0, 1]$) and a major disease with probability x. Given that the patient suffers from a major disease, the physician diagnoses a major disease with probability 1.

4. The patient accepts all minor diagnoses and rejects a major diagnosis with probability $1 - y$ ($y \in [0, 1]$).

5. If the patient accepts the diagnosis, the physician will charge accordingly. Otherwise the patient turns to a second physician and again incurs search costs d. She will then accept any diagnosis with certainty.

The physician's payoff is given as $p_i - c_j$ if the patient accepts the diagnosis, where $i, j \in \{L, H\}$. Otherwise, the physician's payoff is 0. The patient's payoff is given by $B - p_i - nd$. $n \in \{1, 2\}$ indicates the number of physician visits.

The physician's diagnosis is a signal to the patient about her type of disease. When receiving a diagnosis at her first visit, the patient updates her beliefs about her type. When receiving a minor diagnosis, the patient beliefs with probability 1 that she suffers from a minor disease. When receiving a major diagnosis, the patient beliefs to suffer from a minor disease with probability $\frac{(1-h)x}{h+(1-h)x}$. $(1-h)x$ is the probability of having a minor problem and being overcharged. $h + (1-h)x$ reflects the probability of receiving a major diagnosis. In the following two subsections we derive the patients' and the physicians' optimal strategy before turning to the (perfect Bayesian) market equilibria.

3.2.1 Patients' Optimization Problem

The patient maximizes her expected utility by choosing the optimal acceptance rate y of a major diagnosis. Suppose physicians overcharge their patients with probability x. Then the cost function of the patient is given as:

$$
\begin{aligned}
C(y) &= d + (1-h)(1-x)(c_L + e) + [h + (1+\alpha)(1-h)x]yc_H \qquad (3.1)\\
&+ [h + (1-h)x](1-y)\left[d + \frac{(1-h)x(1-x)(c_L+e) + [h + (1+\alpha)(1-h)x^2]c_H}{h + (1-h)x}\right]
\end{aligned}
$$

The patient incurs search costs d for her first visit at a physician. With probability $1-h$ the patient suffers from a minor disease. With probability $1-x$ the physician honestly diagnoses a minor disease. Patients accept a minor diagnosis with certainty because they know that the physicians cannot undertreat. Patients then incur costs of $c_L + e$. With probability $h + (1-h)x$ the patient receives the diagnosis of a major disease. The physician then charges c_H. With probability $(1-h)x$ the patient is overcharged, inducing *trust costs* of $\alpha y c_H$. The patient accepts the major diagnosis with probability y and searches for a second opinion with probability $1-y$. Given that the patient searches for a second opinion she incurs additional search costs d and expected costs of being treated of $(1-h)x(1-x)(c_L+e)+[h+(1+\alpha)(1-h)x^2]c_H$. The patient now accepts both diagnoses with probability 1. With probability $(1-h)x^2$ the patient is defrauded again and therefore again bears *trust costs* of αc_H.

Note that the expected trust costs do not only increase with the overcharging probability x but also with a higher acceptance rate of a major diagnosis y. If the patient never accepts a major diagnosis ($y=0$), she does not face trust costs for the first physician. As patients always accept the second diagnosis, the patient always faces expected trust costs of $\alpha x^2 c_H$ for the second visit of a physician.

The patient trades off between accepting a major diagnosis knowing that she might have been overcharged and costly rejecting the diagnosis. She searches for a second opinion with probability 1 (0) if and only if the costs for a second opinion are smaller (greater) than the expected costs of being overcharged, i. e.

$$d < (>) \left[\frac{(1-h)x}{h+(1-h)x}\right](1-x)[(1+\alpha)c_H - c_L - e] \tag{3.2}$$

In the following, we assume the search costs d to be smaller than $(1+\alpha)c_H - c_L - e$. Otherwise, it is a patient's best response to always accept a major diagnosis from the first physician she visits and to never search for a second opinion.

For the patient to be indifferent between searching for a second opinion and accepting the diagnosis of a major disease both terms have to be equal. Then any $y \in [0,1]$ would be a best response.

Note that equation (3.2) is quadratic in x. Solving for the optimal search strategy y for a given x yields:

Lemma 3.1. *Given a symmetric overcharging strategy by the physicians $x \in [0,1]$, the patients' best response correspondence is given by:*

$$y^*(x) \in \begin{cases} 0 & \text{if } x \in (x_1, x_2) \\ 1 & \text{if } x \in [0, x_1) \cup (x_2, 1] \\ [0,1] & \text{if } x \in \{x_1, x_2\} \end{cases}$$

where

$$\begin{aligned} x_{1/2} = &\frac{1}{2}\left(1 - \frac{d}{[(1+\alpha)c_H - c_L - e]}\right) \\ &\pm \sqrt{\frac{1}{4}\left(1 - \frac{d}{[(1+\alpha)c_H - c_L - e]}\right)^2 - \frac{h}{1-h}\frac{d}{[(1+\alpha)c_H - c_L - e]}} \end{aligned} \tag{3.3}$$

Proof. See *Appendix 3.5.1.* □

The results of *Lemma 3.1* can be seen in *Figure 3.1.* Patients accept a major diagnosis at the first physician if the level of fraud in the market is low (i. e., $x < x_1$). This is because patients anticipate that the physician most probably diagnosed honestly. Patients also accept a major diagnosis if the level of fraud in the market is high ($x > x_2$) as the second physician will most probably diagnose a major disease, too. If $x_1 < x < x_2$ holds, the patient will search for a second opinion.

We now analyze how a change in *trust costs* changes the search behavior of patients.

Proposition 3.1. *An increase in the patients' trust costs ceteris paribus increases the probability of rejecting a major diagnosis from the first physician.*

Proof. See *Appendix 3.5.2.* □

The result of *Proposition 3.1* is displayed in *Figure 3.1.* An increase in the patients' trust costs results in an increase in their expected costs of being defrauded. Thus,

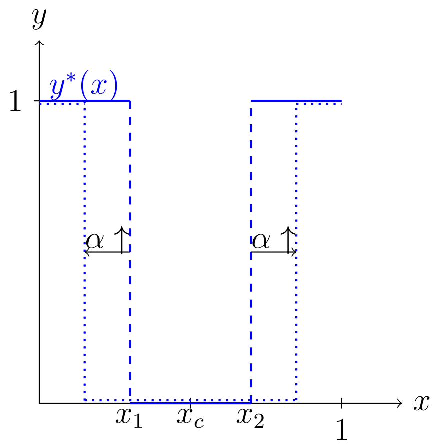


Figure 3.1: Patients' best response correspondence after increase in trust costs.

searching for a second opinion becomes more attractive for patients. In consequence, the probability of rejecting a major diagnosis from the first physician increases.

Note that an increase in the patients' trust costs has the same effect as if the patients were insured and co-payments were increased (Sülzle and Wambach, 2005). Increasing co-payments makes the patients bear a larger share of their treatment costs. This implies that patients also pay a larger share of the fraudulent costs. Thus, the probability of searching for a second opinion increases.

3.2.2 Physicians' Optimization Problem

Physician i maximizes his profit Π_i by choosing the optimal level of overcharging x_i. Due to the given price structure the physician can only make positive profits with patients suffering from a minor disease. Suppose that patients accept a major diagnosis on their first visit with probability y, and all other physicians overcharge the patient with probability x. Then the profit function of each individual physician i yields the following:

$$\max_{x_i} \quad \Pi_i = (1-x_i)e + (1-\beta)x_i \left[\frac{y + x(1-y)}{1 + x(1-y)}\right](c_H - c_L) \tag{3.4}$$

With probability $1 - x_i$, the physician diagnoses patients with a minor disease honestly. An honest diagnosis ensures the physician the certain profit e. With probability x_i, the physician overcharges his patients suffering from a minor disease. $\frac{1}{1+x(1-y)}$ of the patients are on their first visit and accept the major diagnosis with probability y. $\frac{x(1-y)}{1+x(1-y)}$ patients were diagnosed a major disease at another physician, and therefore accept a major diagnosis with certainty. The share of the fraudulent profit $\beta\left[\frac{y+x(1-y)}{1+x(1-y)}\right](c_H - c_L)$ that is being subtracted reflects the *conscience costs* the physician faces when defrauding his patients.

The physician trades off between overcharging patients with a minor disease and treating them honestly. While defrauding yields a higher monetary profit $c_H - c_L > e$ it includes the risk that patients might consult another physician. Thus, the physician's best response is to overcharge any patient suffering from a minor disease with probability 1 (0), if

$$e < (>)(1-\beta)\left[\frac{y+x(1-y)}{1+x(1-y)}\right](c_H - c_L) \tag{3.5}$$

The physician is indifferent between both actions if the mark-up on treating the minor disease equals the expected profit of defrauding. The individual physician's best response is therefore given as:

$$x_i{}^*(x,y) \in \begin{cases} \{0\} & \text{if } e > (1-\beta)\left[\frac{y+x(1-y)}{1+x(1-y)}\right](c_H - c_L), \\ \{1\} & \text{if } e < (1-\beta)\left[\frac{y+x(1-y)}{1+x(1-y)}\right](c_H - c_L) \\ [0,1] & \text{if } e = (1-\beta)\left[\frac{y+x(1-y)}{1+x(1-y)}\right](c_H - c_L), \end{cases}$$

Concentrating on symmetric responses, it has to hold that the physician's individual defrauding strategy x_i corresponds to the other physician's defrauding strategy x. The optimal symmetric defrauding strategy depending on the level of *conscience costs* is given by the following *Lemma*:

Lemma 3.2. *Given the patients' acceptance strategy y, the physicians' symmetric best response for small conscience costs, i. e. $\beta < 1 - \frac{e(2-y)}{c_H - c_L}$, is:*

$$x^*(y) \in \begin{cases} \left\{0, \dfrac{e - y(1-\beta)(c_H - c_L)}{(1-y)[(1-\beta)(c_H - c_L) - e]}, 1\right\} & \text{if } y \in \left[0, \dfrac{e}{(1-\beta)(c_H - c_L)}\right] \\ \{1\} & \text{if } y \in \left[\dfrac{e}{(1-\beta)(c_H - c_L)}, 1\right] \end{cases}$$

For medium conscience costs, i. e. $1 - \frac{e(2-y)}{c_H - c_L} \leq \beta \leq 1 - \frac{e}{c_H - c_L}$, the physicians' symmetric best response is given by:

$$x^*(y) \in \begin{cases} \{0\} & \text{if } y \in \left[0, \dfrac{e}{(1-\beta)(c_H - c_L)}\right] \\ \{1\} & \text{if } y \in \left[\dfrac{e}{(1-\beta)(c_H - c_L)}, 1\right] \end{cases}$$

For large conscience costs, i. e. $\beta > 1 - \frac{e}{c_H - c_L}$, the physicians' symmetric best response is given by:

$$x^*(y) \in \{0\}$$

Proof. See *Appendix 3.5.3.* □

Obviously, for large *conscience costs* the physicians' best response it to always diagnose honestly. For small and medium *conscience costs*, the physicians' best response is a joint best response to the patients' and the other physicians' behavior. Given that patients accept a major diagnosis sufficiently often (i. e., y is large), it is a physician's best response to always defraud patients with a minor disease ($x = 1$). Given that patients rarely accept a major diagnosis, the physician's best response depends on the other physicians' strategy: If the other physicians defraud ($x = 1$), it is the physician's best response to also defraud his patients. This is because he knows that many of his patients are searching for a second opinion and therefore accept any diagnosis with certainty. If the other physicians behave honestly ($x = 0$), it is the physician's best response to also behave honestly. This is because many patients are on their first visit and would search for a second opinion if receiving a major diagnosis. For small *conscience costs* $\beta < 1 - \frac{e(2-y)}{c_H - c_L}$, there also exists a region (the dashed line in *Figure 3.2*) for which physicians are indifferent between overcharging and charging honestly.

Increasing the physician's *conscience costs* ceteris paribus reduces the physician's profit of overcharging $\left(\frac{d\Pi}{d\beta} < 0\right)$. From the physician's individual best response we see that after an increase in *conscience costs* from β_0 to $\beta_1 \leq 1 - \frac{e}{c_H - c_L}$, it is the physician's best response to now even treat a patient with a minor disease honestly if $y \in \left[\frac{e}{(1-\beta_0)(c_H-c_L)}, \frac{e}{(1-\beta_1)(c_H-c_L)}\right]$. If $\beta_1 > 1 - \frac{e}{c_H - c_L}$, it is the physician's best response to now always treat the patient honestly.

Proposition 3.2. *Ceteris paribus, an increase in the physicians' conscience costs decreases their incentive to overcharge.*

Proof. Follows immediately from the physician's individual best response. □

Figure 3.2 illustrates *Proposition 3.2.*

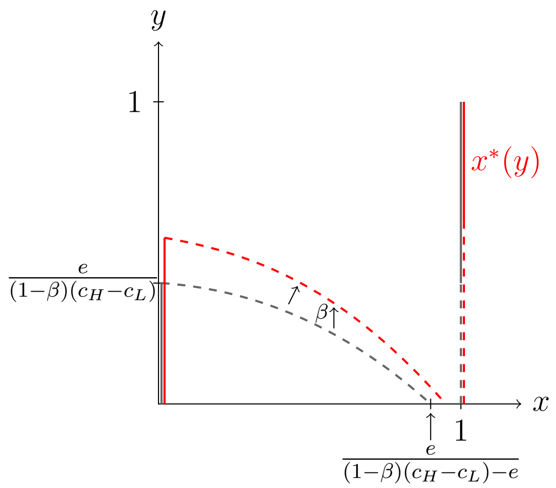


Figure 3.2: Physicians' best response correspondence after increase in conscience costs (for the case of small conscience costs).

3.3 Equilibria of the Model

The market equilibria are obtained by joining the patients' and the physicians' best response. In order to discriminate the effects of introducing patients' *trust costs* and physicians' *conscience costs*, we first consider the benchmark case where patients and physicians only account for monetary incentives, i. e. $\alpha = \beta = 0$.

3.3.1 Equilibria without Other-Regarding Preferences

In a market without other-regarding preferences, Sülzle and Wambach (2005) show that the following *Lemma* holds:

Lemma 3.3. *For $0 < 1 - \frac{e(2-y)}{c_H - c_L}$ and sufficiently small d, there exist three equilibria: In the two mixed strategy equilibria physicians sometimes defraud patients facing a minor disease. Patients sometimes search for a second opinion if they receive a major diagnosis. In the pure strategy equilibrium, physicians always defraud and patients never search for a second opinion.*

If $1 - \frac{e(2-y)}{c_H - c_L} \leq 0 \leq 1 - \frac{e}{c_H - c_L}$, there only exists an equilibrium in pure strategies in which physicians always defraud and patients never search for a second opinion.

If $0 > 1 - \frac{e}{c_H - c_L}$, there exists only an equilibrium in pure strategies in which physicians never defraud and patients never search for a second opinion.

Proof. See Sülzle and Wambach (2005). □

The interesting case is when $0 < 1 - \frac{e}{c_H - c_L}$, i. e. the mark-up for a minor treatment is smaller than the difference in costs between major and minor treatment. Then the intuition for the equilibria is as follows: In the equilibrium of pure strategies (A), patients always accept the diagnosis of a major disease (see *Figure 3.3*). Therefore, it is the physician's best response to always defraud his patients suffering from a minor disease. Anticipating the physicians' behavior, it is a patient's best response to always accept the diagnosis of a major disease because any other physician would defraud the patient, too. In equilibrium B and C, the patients are indifferent between accepting a major diagnosis and costly searching for another opinion while the physicians are indifferent between diagnosing honestly and defrauding a patient with a minor disease. If $0 > 1 - \frac{e}{c_H - c_L}$ it never pays for a physician to defraud a patient. The patient's best response is therefore to always accept a major diagnosis.

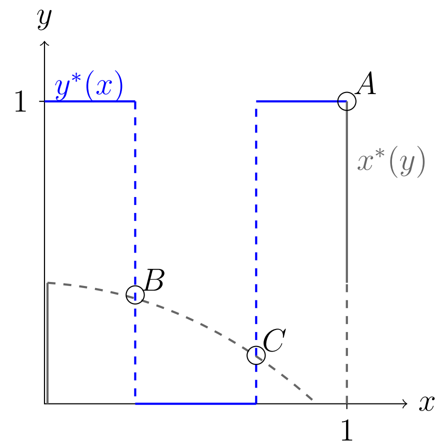


Figure 3.3: Equilibria without other-regarding preferences.

3.3.2 Equilibria with Patients' Trust Costs

Now let us turn to the case where patients lose trust in their physician by a possible overcharging while physicians only care about their monetary profit, i. e. $\alpha > 0$ and $\beta = 0$. A first intuition would lead to the conclusion that an increase of patients' trust costs increases the probability of searching for a second opinion and therefore decreases the physicians' incentive to overcharge. This intuition only partly holds in equilibrium.

Proposition 3.3. *An increase in the patients' trust costs leads to the following changes in the three possible equilibria:*

1. *At equilibrium B', there is less fraud in the market than in B. Patients search less often for a second opinion than in B.*

2. *At equilibrium C', there is more fraud in the market than in C. Patients search more often for a second opinion than in C.*

3. *At equilibrium A, there are no changes.*

Proof. See *Appendix 3.5.4.* □

Proposition 3.3 is illustrated in *Figure 3.4*. Starting from equilibrium, an increase in the probability of searching for a second opinion makes the physician strictly better off if he diagnoses honestly. If a physician is honest, it is the patient's best response to stay at the first physician she visits. In the emerging equilibria, the two effects have to be balanced. Starting from equilibrium B, overcharging in fact decreases and patients search less often for a second opinion than in B. Starting from equilibrium C, the effects are reverse: As more patients search for a second opinion, the physicians anticipate that more patients accept the (second) diagnosis with certainty. Therefore, the physicians' incentive to overcharge increases. Note that both effects – the change in the acceptance rate and the change in the level of fraud – are opposite compared to equilibrium B. Equilibrium A remains unchanged.

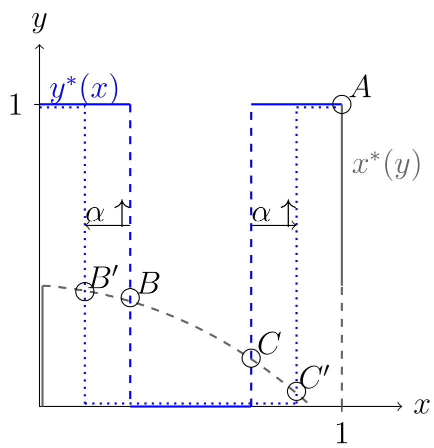


Figure 3.4: Change in equilibria after increase in patients' trust costs.

An increase in the patients' *trust costs* does not only influence the level of overcharging. Through the change in the number of searches for a second opinion, welfare is also affected since each visit at a physician induces search costs of d. Welfare is maximized if patients visit only one physician. Starting from equilibrium B, an increase in *trust costs* decreases the number of second physician visits. Welfare is therefore increased. Again, the opposite holds starting from equilibrium C. Note that overcharging is a pure redistribution between patients and physicians and does therefore not impact welfare.

3.3.3 Equilibria with Physicians' Conscience Costs

In a situation where physicians face *conscience costs* while patients do not account for *trust costs*, i. e. $\alpha = 0$ and $\beta > 0$, the following holds:

Proposition 3.4. *An increase in the physicians' conscience costs does not change the level of fraud in the market as long as conscience costs do not change from small to medium/large or from medium to large. If conscience costs are small, an increase in conscience costs leads to fewer searches for a second opinion.*

Proof. See *Appendix 3.5.5.* □

If *conscience costs* are small, an increase in *conscience costs* does not lead to less overcharging but to a lower search frequency in equilibrium. Fewer searches increase welfare as each search causes costs of d. Equilibrium A does not change. As a consequence, the level of fraud remains unchanged in all three equilibria.

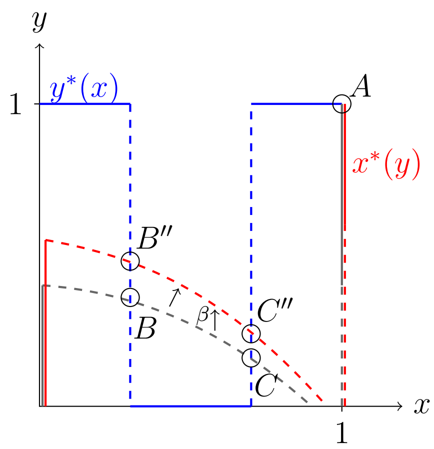


Figure 3.5: Change in equilibria after increase in physicians' conscience costs (for the case of small conscience costs).

The above analysis assumes that the increase in *conscience costs* does not result in a change from small to medium/large or medium to large *conscience costs*. Relaxing this assumption leads to a discontinuity in the physicians' best response at the interval boundaries: An increase from small to medium *conscience costs* then increases the level of fraud in the market if the equilibrium was in mixed strategies (B or C). This is because the two mixed strategy equilibria do not exist when *conscience costs* are medium. Instead, the only market equilibrium is the pure strategy equilibrium in which physicians always defraud and patients always search for a second opinion. An increase from small/medium to large *conscience costs* changes the physicians' behavior to a purely honest behavior because the *conscience costs* then exceed the fraudulent profit.

3.3.4 Equilibria with Patients' Trust Costs and Physicians' Conscience Costs

Accounting for both, patients' *trust costs* and physicians' *conscience costs*, i. e. $\alpha > 0$ and $\beta > 0$, leads to the following results:

Proposition 3.5. *Increasing patients' trust costs and physicians' conscience costs leads to the following changes in the three possible equilibria if conscience costs are small:*

1. *At equilibrium B''', the level of fraud in the market decreases compared to equilibrium B while patients search less often for a second opinion.*

2. *At equilibrium C''', the level of fraud rises compared to equilibrium C. Whether patients search for a second opinion more or less often depends on the cost ratio: If the ratio of trust to conscience costs is high, we observe less searches in C''' than in C while if the ratio is low, we observe more.*

3. *Equilibrium A remains unchanged.*

If conscience costs are medium, physicians always overcharge their patients in equilibrium and patients never search for a second opinion. If conscience costs are *large, physicians always diagnose honestly and patients never search for a second opinion.*

Proof. See *Appendix 3.5.6.* □

The *Proposition* is illustrated in *Figure 3.6.* Starting from equilibrium B, an increase in the patients *trust costs* and the physicians' *conscience costs* leads to fewer searches and therefore to an increase in social welfare. Starting from equilibrium C, increasing the impact of the other-regarding preferences increases welfare if the ratio of *trust costs* to *conscience costs* is high but decreases welfare if the ratio is low.

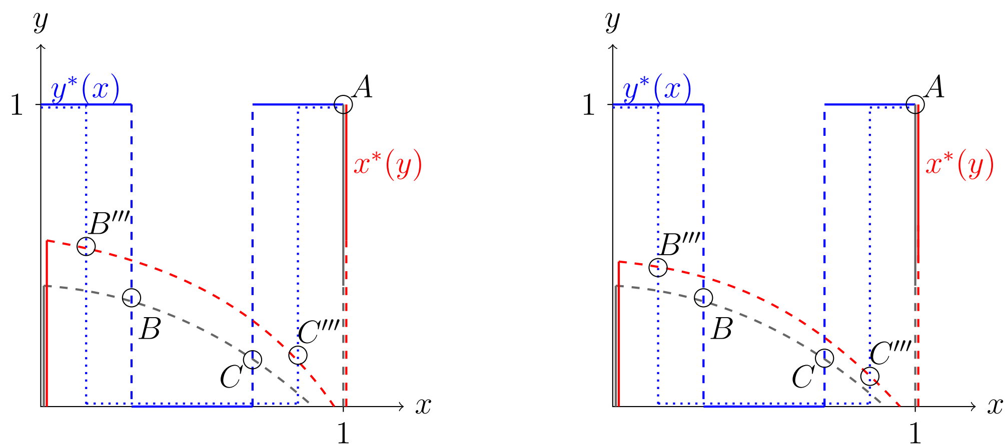


Figure 3.6: Equilibria after increase in *trust* and *conscience costs* given a high (low) cost ratio on the left (right).

3.4 Conclusion

We analyze the impact of an increase in other-regarding preferences on the physicians' level of fraud. In contrast to intuition, we find that an increase in *conscience costs* and *trust costs* does not necessarily decrease the level of fraud in the market, nor does it necessarily decrease the level of second opinions.

Increasing patients' *trust costs* ceteris paribus increases the probability of a search for a second opinion. Starting from equilibrium, increasing the *trust costs* reduces the level of fraud in the market if the level was low before, but increases the level of fraud if the level was high. The increase in fraud is due to the fact that an increase in *trust costs* can lead to more searches for a second opinion in equilibrium, making it more attractive for physicians to defraud their patients. Increasing physicians' *conscience costs* ceteris paribus decreases their incentive to overcharge. Starting from equilibrium, the introduction of *conscience costs* does not change the level of fraud because the higher incentive to diagnose honestly is set off by fewer searches for a second opinion by the patients.

Welfare depends on the number of searches for second opinions and on the ratio of *trust* to *conscience costs*. An increase in the patients' *trust costs* increases (decreases) welfare if the equilibrium is characterized by a low (high) level of fraud. An increase in the physicians' *conscience costs* always increases welfare. Increasing both costs leads to a higher welfare in case of the low fraud equilibrium. In case of the high fraud equilibrium, welfare increases if the ratio of *trust* to *conscience costs* is high.

3.5 Appendix: Proofs

3.5.1 Proof of Lemma 3.1

Suppose $x \in [0,1]$. Rearranging inequality 3.2 (with an equal sign) and substituting with $c = \frac{d}{(1+\alpha)c_H - c_L - e} < 1$ yields:

$$x^2 - (1-c)x + \left(\frac{h}{1-h}\right) c = 0 \tag{3.6}$$

The x_1 and x_2 from *Lemma 3.1* are the (real) solutions to equation (3.6) if

$$(1-c)^2 - 4\left(\frac{h}{1-h}\right) c > 0 \tag{3.7}$$

From equation 3.6 it follows that for $x < x_1$ and $x > x_2$ the patient will strictly prefer to accept a major diagnosis from the first physician, i. e. to choose $y = 1$. If $x_1 < x < x_2$, it is a patients best response to reject the major diagnosis and to search for a second opinion, i. e. to choose $y = 0$. $\square$

3.5.2 Proof of Proposition 3.1

Taking the partial derivative of x with respect to α in equation (3.2) yields:

$$\begin{aligned}
d &= \frac{(1-h)x(\alpha)}{h+(1-h)x(\alpha)}(1-x(\alpha))[(1+\alpha)c_H - c_L - e] \\
\Leftrightarrow 0 &= \left[\frac{(1-h)\frac{dx}{d\alpha}}{h+(1-h)x(\alpha)} - \frac{(1-h)^2 x\frac{dx}{d\alpha}}{(h+(1-h)x(\alpha))^2}\right](1-x(\alpha))[(1+\alpha)c_H - c_L - e] \\
&- \left[\frac{(1-h)x(\alpha)}{h+(1-h)x(\alpha)}\right]\frac{dx}{d\alpha}[(1+\alpha)c_H - c_L - e] + \left[\frac{(1-h)x(\alpha)}{h+(1-h)x(\alpha)}\right](1-x(\alpha))c_H \\
&= \frac{dx}{d\alpha}[1 - \underbrace{\frac{(1-h)x(\alpha)}{h+(1-h)x(\alpha)}\Bigg](1-x(\alpha))[(1+\alpha)c_H - c_L - e]}_{=d} \\
&- \frac{dx}{d\alpha}x(\alpha)[(1+\alpha)c_H - c_L - e] - x^2 c_H + x c_H
\end{aligned}$$

$$\Leftrightarrow \frac{dx}{d\alpha} = -\frac{x(1-x)c_H}{(1-2x)[(1+\alpha)c_H - c_L - e] - d} \tag{3.8}$$

The numerator is always positive, the denominator is positive if $x < x_c$ and negative if $x > x_c$ with $x_c = \frac{1}{2}\left[1 - \frac{d}{(1+\alpha)c_H - c_L - e}\right]$. Given the solutions from *Lemma (3.1)*, it holds that $\frac{dx}{d\alpha} > 0$ for x_2 and $\frac{dx}{d\alpha} < 0$ for x_1. $\square$

3.5.3 Proof of Lemma 3.2

Note that if $\beta > 1 - \frac{e}{c_H - c_L}$ it is obviously the physicians' best response to always charge honestly because the *conscience costs* exceed the fraudulent profit. We therefore assume $\beta < 1 - \frac{e}{c_H - c_L}$ in the following.

Depending on the patients' symmetric strategy $y \in [0, 1]$, we distinguish three situations following Sülzle and Wambach (2005):

1. Patients always accept a major diagnosis from the first physician ($y = 1$): Rearranging inequality (3.5) yields the following:

$$e \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} (1-\beta)(c_H - c_L) \tag{3.9}$$

   It is physician i's best response to always overcharge his patient ($x_i = 1$) because the fraudulent profit $(1-\beta)(c_H - c_L)$ exceeds the profit from an honest diagnosis e.

2. Patients always reject a major diagnosis from the first physician ($y = 0$): Rearranging inequality (3.5) yields the following:

$$e \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} (1-\beta)\left[\frac{x}{1+x}\right](c_H - c_L) \tag{3.10}$$

The physician's best response depends on the other physicians' behavior. We distinguish three cases:

- If all other physicians diagnose honestly, i.e. $x = 0$, equation (3.10) reduces to $e > 0$. Consequently, it is physician i's best response to also diagnose honestly $x_i = 0$. This is because $y = 0$ implies that every patient with a minor disease is on his first visit. If physician i diagnosed a major treatment, the patient would reject the diagnosis with certainty and the physician would make zero profits.
- If all other physicians defraud their patients ($x = 1$), it is physician i's best response to also defraud his patient given that $\beta \leq 1 - \frac{2e}{c_H - c_L}$. Physician i anticipates that there are sufficiently many patients with a minor disease that are on their second visit. They accept a major diagnosis with certainty. If $\beta \geq 1 - \frac{2e}{c_H - c_L}$, $x = 1$ is not a symmetric best response because the deviation strategy $x_i = 0$ would yield higher profits than $x_i = 1$.
- If all other physicians randomize their behavior ($x \in (0, 1)$), a best symmetric response requires physician i to be indifferent between an honest and a fraudulent diagnosis for a patient with a minor disease. From rearranging inequality (3.5) with an equal sign with regard to x we know that indifference holds if and only if

$$x = \frac{e}{(1 - \beta)(c_H - c_L) - e} \tag{3.11}$$

A symmetric best response exists if and only if $\beta < 1 - \frac{2e}{c_H - c_L}$. If $\beta > 1 - \frac{2e}{c_H - c_L}$, an individual physician would again have the incentive to deviate from the equilibrium strategy.

3. Patients do not always but sometimes accept a major diagnosis from the first physician ($y \in (0,1)$). Rearranging inequality (3.5) (with an equal sign) yields:

$$(1-\beta)y(c_H - c_L) + (1-\beta)x(1-y)(c_H - c_L - e) - e = 0 \tag{3.12}$$

Partial differentiation shows that

$$\frac{dx}{dy} = -\frac{(c_H - c_L) - x(c_H - c_L - e)}{(1-y)(c_H - c_L - e)} < 0 \tag{3.13}$$

An increase in y leads to a decrease in x. This is because physicians need fewer patients with a minor disease to be indifferent between defrauding and diagnosing honestly. We again distinguish three cases:

- If the other physicians diagnose honestly ($x = 0$), equation (3.12) is fulfilled if and only if

$$y = \frac{e}{(1-\beta)(c_H - c_L)} \tag{3.14}$$

  Consequently, it is a physician's best response to diagnose honestly ($x_i = 0$) for low enough y $\left(y \in \left[0, \frac{e}{(1-\beta)(c_H - c_L)}\right]\right)$.

- If the other physicians always defraud ($x = 1$), it is a physician's best response to also defraud ($x_i = 1$) if and only if

$$(1-\beta)y(c_H - c_L) + (1-\beta)x(1-y)(c_H - c_L - e) - e > 0 \tag{3.15}$$

  This is given if $\beta < 1 - \frac{e(2-y)}{c_H - c_L}$.

- If the physicians sometimes but not always defraud their patients ($x \in (0,1)$), the individual physician i is indifferent between defrauding and diagnosing honestly if and only if

$$x = \frac{e - y(1-\beta)(c_H - c_L)}{(1 - y[(1-\beta)(c_H - c_L) - e])} \tag{3.16}$$

$x \in (0,1)$ if and only if $\beta < 1 - \frac{e(2-y)}{c_H - c_L}$. According to equation (3.13) x decreases in y and reaches the value 0 for $y = \frac{e}{(1-\beta)(c_H - c_L)}$. $\square$

3.5.4 Proof of Proposition 3.3

As shown in *Proposition 3.1*, an increase in the patients' *trust costs* leads to a change in the patients' best response correspondence. There is no direct influence on the physicians' best response but an indirect influence through the change in the patients' behavior. The change can be obtained by taking the total differential of inequality 3.2 and 3.5 (with an equal sign):

$$\frac{dy}{d\alpha} = -\frac{x(1-x)(1-y)c_H(c_H - c_L - e)}{[(1-2x)[(1+\alpha)c_H - c_L - e] - d][x(c_H - c_L - e) - (c_H - c_L)]} \tag{3.17}$$

The numerator is always positive. The sign of the denominator depends on x. Note that we obtain the same critical value x_c as in *Proposition 3.3*. For $x < x_c$ it holds that the denominator is negative so that $\frac{dy}{d\alpha} > 0$ while for $x > x_c$ the denominator is positive so that $\frac{dy}{d\alpha} < 0$.

3.5.5 Proof of Proposition 3.4

Assume that the increase in *conscience costs* is marginal in the sense that the increase does not switch the costs from small to medium/large or medium to large. Then if the *conscience costs* are small, the emerging equilibria B'' and C'' always exist. This is because $\frac{d}{d\beta}\frac{e}{(1-\beta)(c_H - c_L) - e} > 0$, i. e. the intersection of the physicians' best response and the abscissa is shifted to the right. This in turn ensures the existence of two intersections between the physicians' and the patients' best response.

As seen in *Proposition 3.2*, an increase in the physicians' *conscience costs* ceteris paribus decreases the physicians' level of overcharging. However, the optimal level of overcharging does not change in the two mixed strategy equilibria B'' and C'' as

$$\begin{aligned} x_{1/2} &= \frac{1}{2}\left(1 - \frac{d}{[(1+\alpha)c_H - c_L - e]}\right) \\ &\pm \sqrt{\frac{1}{4}\left(1 - \frac{d}{[(1+\alpha)c_H - c_L - e]}\right)^2 - \frac{h}{1-h}\frac{d}{[(1+\alpha)c_H - c_L - e]}} \\ \Rightarrow \frac{\partial x_{1/2}}{\partial \beta} &= 0 \end{aligned}$$

Note that in the pure strategy equilibrium A, the level of overcharging does not change with increasing *conscience costs*, either.

The increase in *conscience costs* does not have a direct influence on the patients' best response. But there is an indirect influence that increases the patients' acceptance rate y in equilibria B'' and C''. The indirect effect is obtained by taking the total differential of inequality 3.5 (with an equal sign):

$$\begin{aligned} dx &= -\frac{1}{(1-y)^2}dy + \frac{e(c_H - c_L)}{((1-\beta)(c_H - c_L) - e)^2}d\beta = 0 \\ \Rightarrow \frac{dy}{d\beta} &= \frac{(1-y)^2(c_H - c_L)e}{[(1-\beta)(c_H - c_L) - e]^2} > 0 \end{aligned} \tag{3.18}$$

For medium and large *conscience costs* the respective equilibrium in pure strategies still exists.

3.5.6 Proof of Proposition 3.5

Changes in the physicians' behavior follow immediately from *Proposition 3.3* and *3.4*.

Suppose the physicians' *conscience costs* are small. Then the change in the patients' acceptance rate of a major diagnosis is given by:

$$dy = \underbrace{\frac{\partial y}{\partial \alpha}}_{\substack{> 0 \text{ for } x < x_c \\ < 0 \text{ for } x > x_c}} d\alpha + \underbrace{\frac{\partial y}{\partial \beta}}_{>0} d\beta$$

$$= -\frac{x(1-x)(1-y)c_H(c_H - c_L - e)}{[(1-2x)[(1+\alpha)c_H - c_L - e] - d][x(c_H - c_L - e) - (c_H - c_L)]}d\alpha$$
$$+ \frac{(1-y)^2(c_H - c_L)e}{[(1-\beta)(c_H - c_L) - e]^2}d\beta \tag{3.19}$$

For $x < x_c = \frac{1}{2}\left[1 - \frac{d}{(1+\alpha)c_H - c_L - e}\right]$, $\frac{dy}{d\alpha}$ and $\frac{dy}{d\beta}$ are positive. An increase in both costs therefore increases the patients' acceptance rate starting from equilibrium B. For $x > x_c$, $\frac{dy}{d\alpha}$ is negative while $\frac{dy}{d\beta}$ remains positive. The overall effect starting from equilibrium C therefore depends on the cost ratio: Assume α increases for a fixed β. Then the denominator of $\frac{dy}{d\alpha}$ increases, the absolute value of the fraction decreases. We can therefore conclude, that if the ratio of *trust* to *conscience costs* is high, an increase in *trust* and *conscience costs* increases the patients' acceptance rate (i. e. the number of searches for a second opinion decreases). If the ratio is low, the acceptance rate decreases with increasing costs.

If physicians' *conscience costs* are medium or large, an increase in the other-regarding preferences does not change the players' behavior (see *Lemma 3.2*).

4 Reputation in Credence Goods Markets—Experimental Evidence

We experimentally investigate how the intensity of price competition and the possibility to build up reputation influence an expert's incentive to defraud his customers in a market for credence goods. We apply a 2×2 factorial design varying the reputation mechanism (private vs. public histories) and the pricing regime (flexible vs. fixed prices). In contrast to private histories, public histories imply that customers do not only build up their own history but customers observe all customers' histories in terms of undertreatments and prices charged. Our results show that the level of undertreatment and the level of overcharging are significantly lower under fixed than under flexible prices. The level of overcharging is weak significantly higher under public than under private histories if prices are flexible while the opposite holds when prices are fixed. We find the same pattern for the level of undertreatment but differences are not significant.

4.1 Introduction

In this paper, we experimentally investigate an expert's incentive to defraud his customers under different reputation mechanisms and pricing regimes in a credence goods market. In such a market, the customer knows that she suffers from a problem but does not know which type of treatment she needs. The customer must rely on an expert's advice as to which treatment is necessary (Darby and Karni, 1973). The expert's informational advantage may give rise to fraudulent behavior. We look at markets where a customer cannot observe the expert's treatment. Hence, an expert may have an incentive to charge for a more expensive treatment than he actually performed, i. e., the expert may overcharge. Furthermore, the expert may provide a service that is insufficient, i. e., he may undertreat. While customers cannot observe whether they were overcharged even ex-post, customers can verify whether they were undertreated.

Reputation mechanisms may be a feasible instrument to help alleviate expert fraud (Hilger, 2011). Thus, it is important to understand whether and how different reputation mechanisms affect an expert's fraudulent behavior. Previous literature on reputation in credence goods markets has focused on the analysis of a reputation mechanism where each customer can identify the expert she interacts with as opposed to a situation where customers cannot identify experts (Dulleck et al., 2011). If a customer can identify the expert she interacts with, the customer builds up her private history with respect to undertreatment and prices charged. Thus, we will refer to this mechanism as private history mechanism throughout this paper. Dulleck et al. (2011) find that in a competitive market with flexible prices, introducing a private history mechanism neither reduces the level of undertreatment nor the level of overcharging. The contribution of the present paper is twofold: firstly, this work allows for a reputation mechanism that enables customers to observe all customers' histories with respect to undertreatment and prices charged instead of only their own history. We will call this reputation mechanism public history mechanism in the following. Secondly, we allow for reputation building in a market with fixed prices.

Our first contribution is motivated by the recent emergence of public online feedback platforms in credence goods markets. These feedback platforms enable customers to share their history. Intuitively, there are two opposing effects that public histories might have on the expert's incentive to defraud: on the one hand, public histories facilitate customers' reaction to experts' defrauding behavior. If an expert undertreats a customer, under public histories all customers observe that the customer's problem was not fixed. Customers respond to the undertreatment by choosing a different expert in future periods. Thus, the expert undertreating loses most (or all) customers in future periods under public histories. Under private histories, however, the undertreatment is only observed by the customer undertreated. Thus, the expert undertreating may only lose this single customer. While customers can perfectly react to experts who undertreat, a response to experts who overcharge is based on customers' expectation about experts' overcharging behavior. The customers' expectation in turn is based on the prices charged in previous periods. Under public histories, customers observe all prices charged. Hence, customers can base their expectation about an expert's overcharging behavior on more observations than under private histories. Thus, the expert's incentive to defraud his customers should be lower under public than under private histories. On the other hand, if customers can easily switch to experts they expect not to be defrauded by, experts who defraud may try to balance their bad reputation by posting lower prices. If the posted price difference between experts who do and do not defraud is sufficiently large, customers may choose experts offering a low price. A low price in turn makes it less attractive for the expert to treat sufficiently and to charge honestly. Hence, public histories might lead to an increase in the level of fraud. The second effect does not apply if prices are fixed.

The second contribution we make to the literature is to consider reputation mechanisms in markets with fixed prices. Several credence goods markets are characterized by a fixed-price regime, among them the largest credence goods market: the health care market. In Germany, expenditures in the health care market amounted to 264 billion Euro in the year 2008 (Statistisches Bundesamt (Ed.), 2012), of which Transparency International estimates up to 20 billion Euro to be due to fraud (Transparency International - Deutschland e.V. (Ed.), 2008). Under fixed prices, the expert's strategic choices are restricted to whether the expert undertreats and whether he overcharges but there is no possibility to post prices. Thus, experts only compete in two defrauding dimensions but not in terms of prices posted. Hence, customers exclusively choose an expert based on his defrauding history. Thus, we expect a lower level of fraud under fixed than under flexible prices.

A prime example that emphasizes the relevance of our investigation is the *Arztnavigator* ("physician navigator") in Germany. The *Arztnavigator* is a public feedback platform that was established by the three largest statutory health insurance companies in order to assist the insuree in finding a *good* doctor. More than 90% of the general practitioners are listed in the physician directory. As displayed in *Figure 4.4* in the appendix, the *Arztnavigator* polls patients with a standardized questionnaire about their last physician visit. It allows patients to rate their physician in four dimensions: appeal of the practice, quality of communication between the patient and the physician, quality of the treatment, and overall impression. Open comments are not allowed in order to prevent customers from insulting doctors. The feedback information from different patients is pooled for each doctor and made available to other insurees. Insurees can then search for *good* doctors by weighting the results in each feedback dimension. Note that the *Arztnavigator* allows customers to compare the (perceived) quality provided by an expert. In contrast to the implementation of public histories in our experiment, the *Arztnavigator* does not provide any charging information.

We make use of a 2×2 factorial design varying the reputation mechanism between private and public histories as well as the pricing regime between flexible and fixed prices (see *Table 4.1*). We denote the different settings under which the experiment is conducted as "conditions" as opposed to the "treatments" performed by the experts.

The two main results of the paper are: first, under fixed prices the level of undertreatment and the level of overcharging are significantly lower than under flexible prices. Under flexible prices, we observe a price pressure that undermines reputation building: experts do not only compete in the two defrauding dimensions under flexible prices but also in the prices posted. Experts who undertreated in previous periods try to balance their bad reputation by offering low prices. The lower prices make it less attractive for the expert to treat sufficiently and to charge honestly.

Secondly, our results show that the level of overcharging is weak significantly higher under public than under private histories if prices are flexible while the opposite holds when prices are fixed. We find the same pattern for the level of undertreatment but differences are not significant. The additional customer information under public histories intensifies the price competition between experts because experts who defrauded have to balance their reputation not only towards the customer they defrauded but towards all customers. Thus, we observe more fraud under public than under private histories if prices are flexible.

Table 4.1: Conditions.

		Reputation mechanism	
		Private histories	Public histories
Price system	Fixed	PH Fixed	PUH Fixed
	Flexible	PH Flexible	PUH Flexible

The seminal papers on reputation argue that reputation helps reduce inefficiencies in markets with asymmetric information (Kreps and Wilson, 1982; Milgrom and Roberts, 1982). Shapiro (1983) as well as Klein and Leffler (1981) analyze how sellers build up reputation when customers are not able to observe the sellers' product quality prior to purchase. They show that there are price-quality equilibria in which sellers provide high quality at a price above costs. The first contribution on reputational incentives in credence goods markets is by Wolinsky (1993). Experts do not overcharge customers in the first period as customers choose to switch to a different seller in the second period when being charged a major treatment in the first period. This in turn disciplines experts to charge customers honestly. Our experimental results in markets with fixed prices undermine Wolinsky's theoretical argument. We in fact observe that experts build up a reputation by not charging the price for the major treatment. In markets with flexible prices, however, competition drives down the treatment prices. This makes it unprofitable for experts to build up reputation by charging the price for the lower treatment. Instead, the low prices induce a higher overcharging rate.

Ely and Välimäki (2003) and Ely et al. (2008) show that reputation may also lead to more fraud and less efficient market outcomes if experts cannot be held liable for their treatments but treatments are observable. Ely and Välimäki (2003) take up the model by Kreps and Wilson (1982) and consider short-lived customers and two types of long-lived experts: *good* experts who aim to keep their reputation of treating appropriately[34] and *bad* experts who undertreat their customers whenever possible. *Good* experts try to separate themselves from *bad* experts by treating customers such that for the newly entering customers the observed relation of minor to major treatments reflects the ex-ante probabilities of having a minor or a major problem. This implies that *good* experts may undertreat their short-lived customers if there are sufficiently many customers with major problems in previous periods. Note that newly entering customers only observe the type of treatment performed in the previous periods but not whether a treatment was sufficient. In contrast, our long-lived customers do not observe the treatment itself but can verify the treatment's outcome, i. e., customers notice when they are undertreated. Whereas customers in Ely and Välimäki (2003) and Ely et al. (2008) base their strategy on the observed relation of minor to major treatments, customers in our set-up can condition their strategy on whether an expert has undertreated before.

Dulleck et al. (2012) implement a credence goods experiment with fixed prices. The authors are interested in the question of whether *good* experts who always treat sufficiently post high prices or high prices induce a sufficient treatment. They show that in a market without reputational concerns, *good* experts signal their type using the price but high prices do not induce sufficient treatment. In their setting, endogenous prices lead to a more efficient market result. We show that if customers can at least identify the expert they are trading with, i. e., experts can build up a reputation by not undertreating or by charging the price for the minor treatment,

[34] A treatment is appropriate if the type of treatment matches the type of problem, i. e., the expert neither under- nor overtreats.

fixed prices lead to a more efficient market outcome than flexible prices. The reason for the difference in results is that experts compete in an additional dimension in our markets: while in Dulleck et al. (2012) price is the only signal customers can base their beliefs about the experts' treatment behavior on, the reputation mechanisms we implement allow customers to track at least the history of own undertreatments and prices charged. As a consequence, prices become less important as a signaling instrument for sufficient treatment. Instead, experts who undertreated in previous periods balance their bad reputation by posting lower prices.

The remainder of this paper is organized as follows. The next section provides the market description. In section 4.3, we present the experimental set-up including the parameterization. In section 4.4, we identify market equilibria for the given parameterization. Section 4.5 derives predictions. In section 4.6, we present the results. The last section concludes.

4.2 Market

Our model builds upon Dulleck et al. (2011). There are four experts and four customers in the market. We assume that each of the customers either suffers from a minor or a major problem. Each customer knows that she has a problem but does not know which type of problem she suffers from. A customer's ex-ante probability of suffering from a major problem is h, the probability of suffering from a minor problem $1 - h$. These ex-ante probabilities are common knowledge. An expert is able to identify the problem by performing a costless diagnosis.[35] Treating the minor problem costs an expert c_L whereas treating the major problem costs an expert c_H (with $c_H > c_L$). The treatment for the major problem t_H heals both types of problems. The treatment for the minor problem t_L only heals the minor problem. Experts are not liable, i. e., they may treat a customer with a major problem with a minor treatment. The customer cannot observe the treatment but she can verify

[35] We assume zero diagnosis costs in order to make our results comparable to those in Dulleck et al. (2011).

the treatment's outcome, i. e., the customer notices whether the expert undertreated her.[36] Observing undertreatment is feasible because the customer notices whether her problem has been fixed or not (Dulleck and Kerschbamer, 2006, p. 11). The prices for the treatments are denoted by p_L and p_H, respectively.

The stage game depends on the experimental condition. In the following, we outline the stage game for a market with flexible and a market with fixed prices. In both cases, we discriminate between private and public histories by denoting the respective step by $'$ and $''$. The stage game is played repeatedly for n periods for each condition. The stage game for a market with flexible prices is as follows:

1. For each of the customers, nature independently draws the type of problem the customer faces. With probability h a customer suffers from a major problem, with probability $1 - h$ she suffers from a minor problem.

2. Each expert posts a price menu $\{p_L, p_H\}$ for the minor and major treatment.

3.' Each customer observes each expert's price menu posted in the current period and her private history[37] as specified below.

3." Each customer observes each expert's price menu posted in the current period and the public histories as specified below.

4. Each customer chooses an expert or decides not to interact.

5. Each expert observes the type of problem for each customer that chose to interact with him in step 3. Each expert either performs a minor treatment t_L or a major treatment t_H for each of his customer(s).

6. Each expert charges each of his customer(s) the price p_L or the price p_H.

[36] Remember that undertreatment refers to a situation where the customer has a major problem but obtains a minor treatment.

[37] Note that a rational customer builds up her private history in the course of the game and is always aware of her history. Participants in the experiment, however, might forget parts of their history. Therefore, we display the private history in step three of the stage game as a reminder.

7. Each expert observes his payoff and each customer observes her payoff from the current period.

The stage game under fixed prices only differs from the above stage game by not allowing experts to post prices in step two. Instead, the exogenously given prices are common knowledge among the players before the first stage game starts.

The expert's payoff per period and customer π_e is determined by the price p_i he charges less the costs c_j for the treatment t_j he applies ($i, j \in \{L, H\}$) where i and j do not have to coincide.

$$\pi_e = p_i - c_j \qquad\qquad i, j \in \{L, H\} \tag{4.1}$$

If no customer decides to interact with the expert, the expert's payoff amounts to σ. If the customer decides to interact and is not undertreated, the customer derives a utility of v. If she decides to interact and is undertreated, she derives a utility of zero. In either case, the customer has to pay the price for the treatment charged by the expert. The customer's per-period payoff π_c therefore amounts to

$$\pi_c = \begin{cases} v - p_i & \text{if not undertreated}, i \in \{L, H\} \\ -p_i & \text{if undertreated}, i \in \{L, H\} \end{cases} \tag{4.2}$$

if the customer decides to interact. If the customer decides not to enter the market, her payoff amounts to σ.

The information customers observe in step three of the above stage game depends upon the experimental condition:[38]

Private Histories

Under private histories, each customer observes for each of the previous periods the expert she interacted with, the prices posted by this expert, whether this ex-

[38]Note that the categories of information that customers observe are the same as in Dulleck et al. (2011).

pert charged the price for the minor or the major treatment, whether this expert undertreated her, and her profit.

Public Histories

Under public histories, each customer observes for each of the previous periods and each of the customers, the expert the customer interacted with, the prices posted by this expert, whether this expert charged the price for the minor or the major treatment, whether this expert the customer interacted with undertreated her, and what the customer's profit was.

4.3 Experiment

4.3.1 Design

We apply a 2×2 factorial design. In all four conditions, the parameters are fixed and are the same as in the experiment by Dulleck et al. (2011): the ex-ante probability of a customer having a major problem is $h = 0.5$. The expert's costs for providing a minor treatment are $c_L = 2$ and $c_H = 6$ for a major treatment. The customer derives a utility of $v = 10$ if her problem is solved. Otherwise, the customer's utility amounts to $v = 0$. In case no interaction takes place, customers and experts receive a payoff of $\sigma = 1.6$ (outside option).

The stage game is repeated for 16 periods. In all conditions, we use matching groups of eight players. The assignment of the eight players to a matching group does not change throughout the experiment. Four of the players take the role of a customer. The remaining four take the role of an expert. The roles are randomly assigned at the beginning of the experiment and do not change throughout the 16 periods. Across conditions, we vary the reputation mechanism between private and public histories and the pricing regime between fixed and flexible prices.

In the flexible-price conditions, experts announce prices $\{\{p_L, p_H\} \in \mathbb{N}^2 | 1 \leq p_L, p_H \leq 11, p_L \leq p_H\}$ in step two of the stage game. In the fixed-price conditions, we set the exogenously given prices $\{p_L, p_H\} = \{4, 8\}$ in periods 1–9, and

$\{p_L, p_H\} = \{0, 3\}$ in periods 10–16. In periods 1–9, there is no obvious way to choose the fixed prices. We use the price vector of $\{p_L, p_H\} = \{4, 8\}$ for three reasons: firstly, equal mark-up prices are observed in several credence goods markets with price regulation. An example is the remuneration of short counseling interviews at general practitioners in Germany. Although the content of the counseling interviews and thus the service usually varies widely, the remuneration for the physician is always the same. Assuming that the general practitioner has similar costs for a short counseling independent of the content, he faces equal mark-ups for different services. Another example is the market for taxi rides. Once the passenger has paid the flat fee for taking a taxi, the taxi driver earns the same mark-up for each additional kilometer he transports the passenger (assuming that traffic is similar). The second reason we implement the price vector $\{4, 8\}$ is that the two equal mark-up vectors $\{4, 8\}$ and $\{3, 7\}$ are the most frequently posted price vectors under flexible prices in our experiment. Thus, by choosing one of the two equal mark-up vectors, we approximate the expert pricing behavior observed under flexible prices. Among the two most often posted price vectors, we thirdly choose to implement the vector $\{4, 8\}$ as this is one of the vectors used in the related study by Dulleck et al. (2012) that we will compare our results to.[39]

In periods 10–16, the price for the major treatment $p_H = 3$ is derived from the predicted expert pricing behavior.[40] The level of p_H ensures that customers still interact although they expect to be undertreated and overcharged in equilibrium. Theory does not provide a prediction for the price p_L as it is never charged in equilibrium. As we implement equal mark-up prices in the first nine periods, we approximate the equal mark-up price by setting p_L to the minimum of $p_L = 0$ in

[39]Under flexible prices, theory predicts that experts post a price p_H that is below marginal costs for the major treatment. Thus, experts make losses if they do not undertreat a customer with a major problem. Inducing expert losses exogenously by setting a price that is below costs for the major treatment may increase experts' undertreatment compared to a situation where the price choice is endogenous. Thus, we fix the price above the costs for a major treatment.

[40]Note that a price below marginal costs does not alter the experts' incentive to provide a sufficient treatment in the last periods because experts undertreat independent of the price vector posted.

periods 10–16. Note that experts under flexible prices also posted a price for the minor treatment in periods 10–16 that was on average slightly below costs.

4.3.2 Procedure

The experimental sessions were conducted in the Cologne Laboratory for Economic Research between March and November 2012. 256 participants took part in the experiment. Participants were equally allocated to the four conditions so that in each condition there were 64 participants. Hence, there were eight matching groups (markets) per condition. We used ORSEE (Greiner, 2004) to recruit participants. We ran the experiments using z-Tree (Fischbacher, 2007). None of the participants took part in more than one session. The instructions were read aloud at the beginning of each session. A detailed set of control questions followed the instructions in order to ensure that all participants understood the experiment. After the experiment, players' social preferences were determined by the choice of payoff pairs for oneself and a randomly assigned other person. Additionally, we used a questionnaire to control for gender and age. The average time each session lasted was two hours. Participants earned on average 20.07 Euro.

4.4 Equilibria of the Model

In the following, we look for perfect Bayesian equilibria of the game described in section 4.2.[41] Two classes of equilibria might emerge: *no-reputation equilibria* and *reputation equilibria* (see Dulleck et al., 2011). The no-reputation equilibria are the one-shot Bayesian equilibria that are played repeatedly over all 16 periods while the reputation equilibria are based on the players' repeated interaction. In what follows, we will characterize both classes of equilibria for the flexible- and fixed-price conditions.

[41]Note that the outlined equilibria are not exhaustive. There exist, for example, also equilibria with asymmetric expert behavior as pointed out by Dulleck et al. (2011). In line with their analysis, we restrict our analysis to equilibria with symmetric expert behavior.

4.4.1 Flexible Prices

The equilibria under flexible prices are adapted and generalized from Dulleck et al. (2011). We do not distinguish between private and public histories because for both classes, there is no difference in the classes between the two types of histories.

Lemma 4.1. *No-reputation equilibria*
In a market with flexible prices, there exist equilibria with the following characteristics: all experts post a price menu $\{n.d., 3\}$[42] *with probability* $x = 0.84398$ *and an unattractive price menu* $\{n.d., p_H\}$ *where* $p_H > 3$ *with probability* $1 - x$*. If an expert posts* $\{n.d., 3\}$*, the expert undertreats customers facing a major problem and always overcharges his customers. If no expert posts* $\{n.d., 3\}$*, there is no interaction.*

Proof. See Dulleck et al. (2011). □

In a market with flexible prices, there exist the above described no-reputation equilibria in which experts strongly compete in the price dimension. The competitive price $p_H = 3$ for the major treatment is so low that experts would even make losses in expectation if they always overcharged but did not undertreat customers with a major problem. Thus, it does not pay for experts to build up reputation. Hence, experts always undertreat customers with a major problem and always overcharge.

Experts are indifferent between posting a competitive price vector $\{n.d., 3\}$ with probability $x = 0.84398$ and posting an unattractive price vector where $p_H > 3$ with probability $1 - x$. The reason experts play mixed strategies is that in case an expert only serves one customer, his profit amounts to 1 while the outside option yields a payoff of 1.6. Thus, it only pays for the expert to offer the competitive price if he can expect more than one customer on average.

We next turn to the class of reputation equilibria which is characterized as follows:

Lemma 4.2. *Reputation equilibria*
In a market with flexible prices, there exist equilibria with the following characteristics: each expert posts a price menu $\{n.d, 5\}$ *in the first nine periods and a price*

[42] n. d. denotes 'not determined'.

menu $\{n.d., 3\}$ *afterwards. Experts always overcharge customers with a minor problem. Experts do not undertreat customers facing a major problem in the first nine periods with sufficiently high probability. Experts undertreat customers with a major problem in periods* 10–16 *and overcharge all customers.*

Proof. See *Appendix 4.8.1.* □

In contrast to the no-reputation equilibrium, the reputation equilibrium implies that experts post higher prices in the first periods allowing them to build up a reputation by not undertreating without making losses. A customer observes her own history (under private histories) respectively all customers' histories (under public histories) regarding undertreatment and prices charged. The customer conditions her choice of expert on whether she respectively any of the four customers has been undertreated by the expert in previous periods. In periods 10–16, experts undertreat customers with a major problem because the additional current payoff from undertreating is larger than the forgone profit when losing the customers in the last periods. Experts always overcharge customers with a minor problem in equilibrium. In periods 10–16, experts also overcharge the customers facing a major problem. Customers still interact in periods 10–16 because the price for the major treatment is sufficiently low. Thus, the expected payoff from interacting exceeds the outside option.

The reasoning why experts only build up reputation in the treatment but not the charging dimension under flexible prices is as follows: customers cannot observe whether an expert overcharged. Thus, customers' strategy can only condition on whether the price for the minor or the major treatment was charged. If an expert tried to build up reputation of not overcharging by always charging the price for the minor treatment (with $p_L < p_H = 5$) in the first periods, his payoff would be lower than the outside option. In later periods, the expert who built up reputation by charging the price for the minor treatment would have to return to charging the price for the major treatment in order to make positive profits. Due to price competition, the competitors would slightly undercut the higher price in later periods though and still offer a sufficient treatment. Hence, the competitors would attract all customers. Thus, charging p_L in the first periods is not profitable under flexible prices.

4.4.2 Fixed Prices

In contrast to flexible prices, the experts' action space reduces to the treatment and charging choice if prices are fixed. In the following, we present the classes of equilibria that have a similar structure as the equilibria in a market with flexible prices. Under fixed prices, there exist no-reputation equilibria with the following properties:

Lemma 4.3. *No-reputation equilibria*

In a market with fixed prices, there exist equilibria in which there is no interaction in periods 1–9. *In periods* 10–16*, experts always overcharge customers and undertreat those customers with a major problem.*

Proof. See *Appendix 4.8.2.* □

In contrast to the no-reputation equilibria under flexible prices, prices are not low enough for customers to interact in periods 1–9. Their outside option of 1.6 is larger than the expected payoff of interacting which amounts to -3 given that experts always overcharge and undertreat customers with a major problem. Thus, customers do not interact in the first nine periods. In the last periods, prices in the fixed price set-up are low enough so that although experts always overcharge and always undertreat customers with a major problem, customers interact.

With prices given by $\{p_L, p_H\} = \{4, 8\}$ in periods 1–9 and by $\{p_L, p_H\} = \{0, 3\}$ in periods 10–16, customers would not interact in the first nine periods if they randomized between experts. If interaction is still observed, customers must hence coordinate on experts. The class of reputation equilibria in a market with fixed prices is characterized by:

Lemma 4.4. *Reputation equilibria*

In a market with fixed prices, there exist equilibria in which experts do not undertreat *customers with a major problem in periods* 1–9 *but always overcharge customers with a minor problem. In periods* 10–16*, experts always overcharge and undertreat customers with a major problem.*

Proof. See *Appendix 4.8.3.* □

The reputation equilibria outlined above are characterized by experts building up a reputation by not undertreating in periods 1–9. Experts always undertreat in periods 10–16. Experts overcharge customers with a minor problem in all periods. The equilibria described in *Lemma 4.4* hold for both information structures: private and public histories.

Under public histories, there exists an additional reputation equilibrium in which experts build up reputation in the first periods by not undertreating and always charging p_L in periods 1–7. The expert serving the customers makes zero profits in the first periods. In periods 8 and 9, the expert makes positive profits by charging the customers the major treatment and not undertreating. In periods 10–16, experts always overcharge customers and undertreat those customers facing a major problem.[43] The equilibrium is characterized as follows:

Lemma 4.5. *Reputation equilibrium without overcharging*
In a market with fixed prices and public histories, there exists an equilibrium in which experts do not undertreat in periods 1–9*. In periods* 10–16*, experts always undertreat customers with a major problem. In periods* 1–7*, experts charge* p_L*; in periods* 8–16*, experts always overcharge customers with a minor problem.*

Proof. See *Appendix 4.8.4.* □

If the expert serving customers deviated in periods 1–7 by undertreating a customer, by charging a customer p_H, or both, all customers would observe the deviation. Consequently, all customers would visit a different expert. This in turn

[43] As outlined before, a reputation equilibrium without overcharging does not exist under flexible prices. This is because competitors would slightly undercut the price of the expert not overcharging in periods 8 and 9.

disciplines the experts not to deviate. Note that the punishment mechanism for charging p_H only works under public but not under private histories. If a customer under private histories observed that he was charged p_H, visiting a different expert would not be a credible threat because being the only customer at the other expert would mean that she would be undertreated and overcharged, yielding a payoff of -3.[44] If the customer did not interact after the deviation instead of switching to a different expert, her payoff amounted to 1.6 while staying with the expert charging p_H yields a payoff of 2. Thus, customers charged p_H would still visit the same expert after being charged p_H. Thus, the deviation is profitable for experts under private histories.

4.5 Predictions

In the following, we derive predictions for the differences in the level of undertreatment and the level of overcharging between the four conditions. We also shortly describe experts' price posting behavior.

4.5.1 Level of Undertreatment

The first hypothesis is concerned with the difference between the flexible- and fixed-price conditions. There is no trade in the first nine periods under fixed prices if players coordinate on the no-reputation equilibrium. Thus, if we observe interaction, theory predicts that experts and customers behave according to the reputation equilibrium. Then, none of the customers is undertreated in periods 1–9. Under flexible prices and observed trade, however, players might either coordinate on the reputation equilibrium or the no-reputation equilibrium. Hence, undertreatment is possible in periods 1–9. Independent of which equilibrium players coordinate on under flexible prices, the equilibrium price for the major treatment in the flexible-price

[44] Remember that under private histories customers do not observe how many customers an expert had in the previous period.

condition is lower than the exogenously set price in the fixed-price conditions. Thus, we can state the following hypothesis:

Hypothesis 4.1 (Fixed vs. flexible prices: undertreatment)**.** *If interaction between experts and customers is observed in periods* 1–9*, the level of undertreatment in periods* 1–9 *is equal or lower under fixed than under flexible prices. The price p_H experts post under flexible prices is lower than the exogenously set price p_H in the fixed-price conditions.*

Next, we turn to the difference in the level of undertreatment if prices are flexible and histories are either private or public. Whether the level of undertreatment is lower or higher under private than under public histories depends on whether experts and customers coordinate on the no-reputation or reputation equilibrium. Thus, theory does not provide a prediction about the treatment comparison. The intuition why we expect less undertreatment under public than under private histories is that each customer observes the sufficiency of all treatments provided and not only the sufficiency of the her own treatment. Hence, the customer has the possibility to condition the choice of expert on whether any of the customers has been undertreated by the expert in any of the previous periods under public histories. Therefore, expert punishment and the customers' search for an expert not undertreating is facilitated. Prices, however, do not differ between the two reputation mechanisms. Thus, we can derive the following hypothesis:

Hypothesis 4.2 (Private vs. public histories under flexible prices: undertreatment)**.** *If players play a reputation equilibrium in a market with flexible prices, we expect the level of undertreatment in periods* 1–9 *to be equal or lower for public than for private histories. The price p_H posted by experts does not differ between the conditions.*

The above outlined reasoning for why we expect less fraud under public than under private histories also applies to a market with fixed prices. We still distinguish in the hypotheses for the difference between private and public histories with respect to the price regime because results differ across price regimes as shown below.

Hypothesis 4.3 (Private vs. public histories under fixed prices: undertreatment)**.** *If experts and customers play the reputation equilibrium outlined in Lemma 4.4 or in Lemma 4.5 in a market with fixed prices, we expect the level of undertreatment in periods* 1–9 *to be equal or lower for public than for private histories.*

4.5.2 Level of Overcharging

Next, we turn to the hypotheses concerning overcharging. Under flexible prices, customers are always overcharged. If experts charged p_L instead of p_H in the first periods, experts would have lower payoffs in expectation than if not interacting. Thus, charging p_L might only be rational if higher prices in later periods compensated the forgone profit in the first periods. Due to price competition, the competitors would undercut the higher price in later periods though and still offer a sufficient treatment. Thus, charging p_L in the first periods is not profitable under flexible prices. Under fixed prices and public histories, however, there exists a reputation equilibrium in which customers are not overcharged in the first seven periods. Therefore, we state the following hypothesis with respect to the expected difference in overcharging in the two price regimes:

Hypothesis 4.4 (Fixed vs. flexible prices: overcharging)**.** *If interaction between experts and customers is observed in periods* 1–9*, the level of overcharging in periods* 1–9 *is equal or lower in a market with fixed prices than in a market with flexible prices.*

Under flexible prices, experts cannot build up reputation by treating sufficiently and charging p_L. This insight holds independent of the information structure in the market. The reason is that experts' possibility to undercut competitors' prices is sufficient for the non-existence of no-overcharging equilibria in the markets we consider. Thus, we expect no difference in the level of overcharging between public and private histories:

Hypothesis 4.5 (Private vs. public histories under flexible prices: overcharging)**.** *If interaction between experts and customers is observed in periods* 1–9*, the level of* *overcharging in periods* 1–9 *does not differ between public and private histories in a market with flexible prices.*

In a market with fixed prices and public histories, experts can build up reputation with respect to the sufficiency of a treatment and the charging decision. The equilibrium outlined in *Lemma 4.5* shows that experts charge customers p_L in periods 1–7. Customers can credibly threaten experts to switch to a different expert if the expert undertreats or charges p_H. This is because all customers observe an expert's deviation. Losing all customers induces a sufficiently high reduction in expert profits such that experts will not charge p_H in the first periods. Under private histories, however, customers cannot credibly threaten to switch to a different expert because other customers would not observe the deviation and thus would not punish the expert.[45] Visiting an expert who only serves one customer is not rational for the customer as she would be undertreated. Thus, visiting a different expert is not a credible threat. Hence, we can state the following prediction with respect to the difference between private and public histories under fixed prices:

Hypothesis 4.6 (Private vs. public histories under fixed prices: overcharging)**.** *If interaction between experts and customers is observed in periods* 1–9*, the level of overcharging in periods* 1–9 *is equal or lower under public than under private histories in a market with fixed prices.*

4.6 Results

This section provides an overview and a discussion of the methods used before presenting the experimental results for the level of undertreatment and the level of overcharging. We also shortly comment on the prices posted and charged. We restrict our analysis to the first nine periods where reputational concerns may play a role.

[45] Note that under private histories, customers cannot observe how many other customers visit an expert.

All non-parametric test results reported in the following are based on two-tailed Mann Whitney U Tests. Test results are reported to be (weakly) significant if the two-tailed test's p-value is smaller than 0.05 (0.1). We consider the average per market over individuals and over the first nine periods as one independent observation. Thus, our non-parametric test results are based on eight independent observations per condition.

In order to separate the mechanisms at work and to account for individual heterogeneity, we complement the non-parametric test results by parametric tests in form of regressions. The data structure, however, is challenging for regression analysis as several factors combine: firstly, the stage game is repeated for 16 periods which imposes a serial correlation between the observations per individual over time. Secondly, eight individuals interact within a market which potentially leads to correlated observations within the market. And thirdly, our dependent variables—whether a customer was undertreated respectively overcharged in a period—are binary.

We follow Dulleck et al. (2011) and make use of the random effects panel probit regression with standard errors clustered on the individual level. The panel probit model accounts for the fact that the stage game is repeated for 16 periods and that the dependent variable is binary. In contrast to the fixed effects estimator, the random effects estimator allows us to estimate the treatment effect although the condition does not vary within an individual. Note that current implementations of binary panel regressions do not allow clustering on a different level than the individuals' level nor is it possible to use robust standard errors. Thus, we may not be capturing the correlation within markets. Introducing market dummies to control for the different markets is not an option firstly because the dummies introduce high collinearity and secondly because results would be relative to the reference market.

Therefore, we also present the results of a random effects panel OLS regression with robust standard errors clustered on the market level (see for the methodology of robust clustered standard errors Huber, 1967; Rogers, 1993; White, 1980). This alternative approach has been previously used by Dulleck et al. (2012) in the same set-up as ours. In contrast to the implementation of the panel probit regression, the implementation of the panel OLS regression does explicitly allow to cluster standard errors on the market level. There are two more advantages of the panel OLS results: firstly, the panel OLS regression eases the interpretation of coefficients. Secondly, the interaction term cannot be misleading (see for a methodological discussion on the interpretation of interaction terms in non-linear response models Ai and Norton, 2003). The drawback of the panel OLS regression is that it does not account for the binarity of the dependent variable and hence suffers from out-of-bound predictions and built-in heterogeneity (Wooldridge, 2009).

Our main results hold independent of the choice of method. Whenever the panel OLS estimates deviate from the panel probit estimates, we indicate the deviation in a footnote.

In order to evaluate the model fit, we report the McKelvey and Zavoina $R^2_{M\&Z}$ (McKelvey and Zavoina, 1975) for the binary panel models and the ordinary R^2 for the panel OLS regressions.[46] We choose the McKelvey and Zavoina $R^2_{M\&Z}$ because it is applicable to binary panel models and mirrors the ordinary R^2. It is defined as $R^2_{M\&Z} = Var(\hat{y}^*)/(Var(\hat{y}^*) + Var(\epsilon))$ where $\hat{y}^*$ denotes the prediction of the latent response variable and ϵ denotes the error term. By assumption of the probit model $Var(\epsilon) = 1$. Note that due to this assumption about the variance of the error term, the ordinary R^2 and the McKelvey and Zavoina $R^2_{M\&Z}$ are not identical.

4.6.1 Level of Undertreatment

The descriptive experimental results for the level of undertreatment are presented in *Table 4.2*. As our design exactly builds upon the design of Dulleck et al. (2011), we present their corresponding results. The additional data allows us to compare the level of undertreatment under private and public histories with a situation in which the customer cannot even identify the expert she is interacting with (condition *None*), i. e., a market without reputational concerns.

[46]Note that due to the different estimation technique, we cannot compare probit and OLS models with the same R^2.

Table 4.2: Percentage of undertreatment in periods 1–9.

		Reputation mechanism			
		This paper		Dulleck et al. (2011)	
		Private histories	Public histories	None	Private histories
Price system	Fixed	31.43%	24.41%	–	–
	Flexible	58.47%	63.46%	61.18%	59.22%

In our regressions on the level of undertreatment, we control for the period in which an interaction takes place, the conditions and the interaction effect between the conditions. The basic specification is as follows where c_i denotes the random intercept of individual i and u_{it} denotes the idiosyncratic error term for individual i in period t:

$$\begin{aligned} undertreatment_{it} = \beta_0 + \beta_1 period_{it} + \beta_2 private\_histories_{it} + \beta_3 fixed\_prices_{it} \\ + \beta_4 private\_histories_{it} \cdot fixed\_prices_{it} + c_i + u_{it}. \end{aligned} \tag{4.3}$$

Table 4.3 displays our regression results. We report the random effects panel OLS estimation with robust standard errors clustered on the market level in the last two columns.

Result 4.1 (Flexible vs. fixed prices: undertreatment)**.** *The level of undertreatment is significantly higher under flexible than under fixed prices. Prices posted by experts under flexible prices are significantly lower than the exogenously given prices in the fixed-price condition.*

Our experimental results are in line with our first hypothesis: the level of undertreatment is significantly higher in the flexible- than in the fixed-price regime (see models (3) and (5) in *Table 4.3*; Mann Whitney U Test: $p < 0.001$). According to the OLS estimates, this difference in the level of undertreatment amounts to 33.3 percentage points. Prices posted by experts under flexible prices are significantly

Table 4.3: Random effects panel regressions on undertreatment in periods 1–9.

	Panel probit				Panel OLS	
Undertreatment	(1)	(2)	(3)	(4)	(5)	(6)
Period	0.046*	0.047*	0.044*	0.044*	0.015	0.015
	(0.027)	(0.027)	(0.026)	(0.026)	(0.013)	(0.013)
Private histories		0.134	0.068	−0.133	−0.020	−0.043
		(0.187)	(0.155)	(0.216)	(0.071)	(0.100)
Fixed prices			−0.955***	−1.161***	−0.333***	−0.393***
			(0.160)	(0.227)	(0.072)	(0.114)
Private histories ·				0.415		0.125
fixed prices				(0.312)		(0.141)
Intercept	−0.375**	−0.446**	0.064	0.171	0.531***	0.564***
	(0.161)	(0.190)	(0.187)	(0.204)	(0.093)	(0.098)
$R^2_{M\&Z}$	0.029	0.033	0.184	0.190	–	–
R^2	–	–	–	–	0.117	0.120
Observations	454	454	454	454	454	454

Standard errors are clustered on the individual level for panel probit regressions (Note: clustering for panel probit regressions on a different level than the individuals' is not implemented). Standard errors are robust and clustered on the market level for panel OLS regressions. Standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. p-values are based on two-tailed tests.

lower than the exogenously given prices in the fixed-price condition (Mann Whitney U Test: $p < 0.001$ for both treatment prices).[47]

Players coordinate on the reputation equilibrium under fixed prices. Experts build up a reputation by treating customers sufficiently in the first periods. The average rate of undertreatment under fixed prices amounts to 27.59% in the first nine periods (and even only to 22.97% in the first eight periods). Customers punish experts who undertreat more often under fixed than under flexible prices by returning significantly less often to the undertreating expert (Mann Whitney U Test: $p < 0.001$). In later periods, experts exploit their reputation by undertreating. The average rate of undertreatment rises to 86.70% in periods 10–16 under fixed prices. *Figure 4.1*

[47] Note that this difference in the level of undertreatment holds for both types of reputation mechanisms: private (Mann Whitney U Test: $p = 0.009$) and public histories (Mann Whitney U Test: $p = 0.006$).

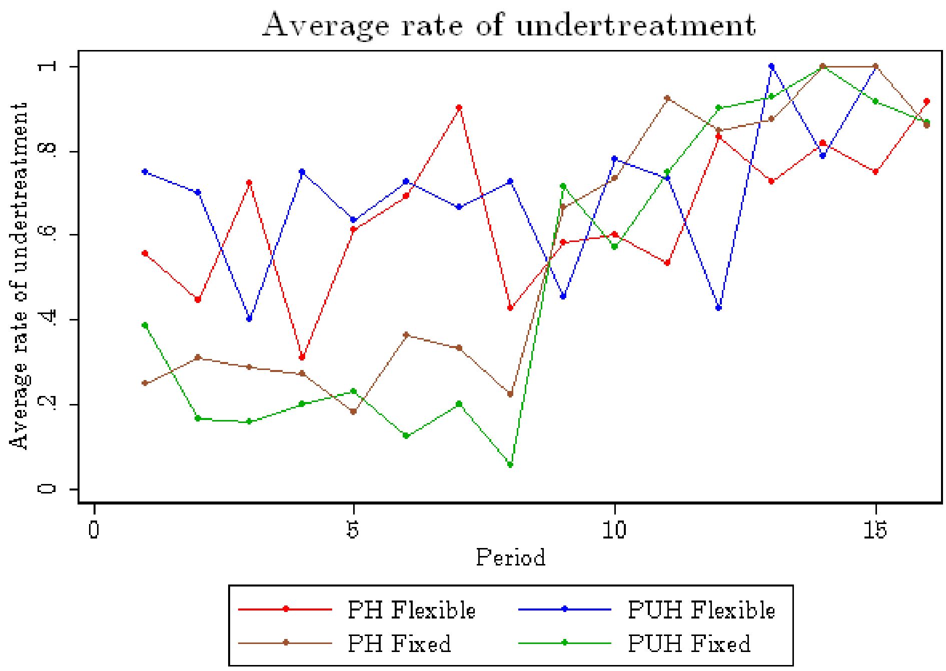


Figure 4.1: Average rate of undertreatment for each condition.

illustrates the average rate of undertreatment for each of the four conditions over time.

Under flexible prices, players rather coordinate on the competitive than the reputation equilibrium. The average rate of undertreatment amounts to 60.81% in the first nine periods. We observe a price pressure that undermines reputation building: experts do not only compete in the two defrauding dimensions under flexible prices but also in the prices posted. *Figure 4.2* visualizes the decline in the price posted for the major treatment over time in the two flexible-price conditions. Experts who undertreated in previous periods try to balance their bad reputation by offering low prices in the following periods. In fact, we find that the average price posted for the major treatment prior to an expert's first undertreatment amounts to 7.30 while the price significantly declines to 6.04 on average after an expert's first undertreatment (Mann Whitney U Test: $p < 0.001$). Concerning the customer behavior under flexible prices, our results show that customers return significantly more often to an expert that has undertreated the customer in one of the previous periods under the

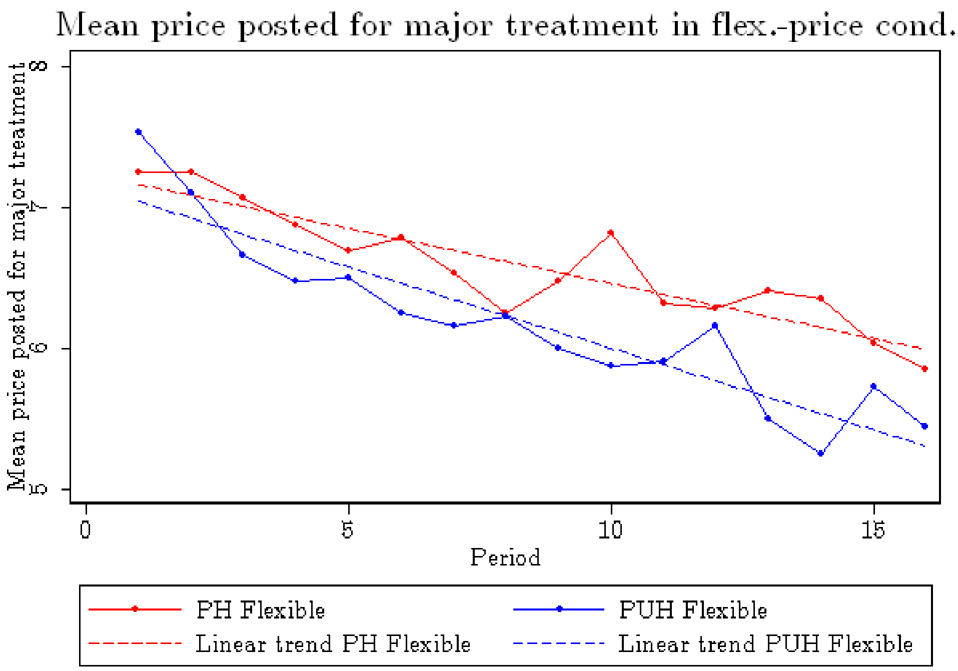


Figure 4.2: Development of prices posted over time in flexible-price conditions.

flexible- compared to the fixed-price conditions (Mann Whitney U Test: $p < 0.001$). Thus, the decline in the price posted for the major treatment by an expert who undertreated seems to be sufficiently large to attract customers in future periods. Hence, price competition undermines reputation building and thereby leads to a higher level of undertreatment if prices are flexible.

For both pricing regimes—fixed as well as flexible—experts' choice of treatment in periods 10–16 matches the predicted behavior. Following the theoretical predictions, we would expect experts to always undertreat in the last seven periods. The average level of undertreatment under flexible prices in periods 10–16 amounts to 77.78% and under fixed prices—as mentioned above—even to 86.70%.

Note that our results are in contrast to Dulleck et al. (2012): Dulleck et al. (2012) find in a setting without reputation that flexible prices lead to a more efficient market outcome than fixed prices because experts signal their type using a high price for the major treatment. We find that under reputation, flexible prices increase the level of undertreatment (see *Table 4.2* and *4.3*) and decrease market efficiency com-

Table 4.4: Efficiency in periods 1-9.

		Reputation mechanism	
		Private histories	Public histories
Price system	Fixed	60.89%	84.88%
	Flexible	58.55%	64.95%

Efficiency is calculated as follows: (actual average surplus − outside option)/(maximum average surplus − outside option).

pared to a market set-up with fixed prices (see *Table 4.4*; Mann Whitney U Test: $p = 0.036$).[48] The difference between Dulleck et al. (2012) and our results can be explained by the additional dimension of competition that experts face in a market where reputation mechanisms are in place: while in Dulleck et al. (2012) price is the only signal customers can base their beliefs about the experts' treatment behavior on, the reputation mechanisms we implement allow customers to track at least the history of own undertreatments. Thus, customers can base their decision on which expert to visit not only on the price but also on the sufficiency of treatments in previous periods. As a consequence, prices become less important as a signaling instrument. Instead, experts use the price to balance their reputation once they undertreated. Hence, the additional dimension of expert competition changes experts' pricing and defrauding behavior.

Result 4.2 (Private vs. public histories under flexible prices: undertreatment)**.** *The level of undertreatment is non-significantly higher under public than under private histories if prices are flexible. Prices are non-significantly lower under public than under private histories if prices are flexible.*

In contrast to *Hypothesis 4.2*, we find that more customer information does not lead to a decrease in the expert's incentive to undertreat his customer if prices are

[48] Note that the efficiency result is driven by the difference in the public history conditions (Mann Whitney U Test: $p = 0.010$). In the private history conditions, the difference between fixed and flexible prices is not statistically significant (Mann Whitney U Test: $p = 0.563$). The lower efficiency level under private histories with fixed prices is due to a low rate of interaction. Several customers in this condition decided not to interact for almost all 16 periods.

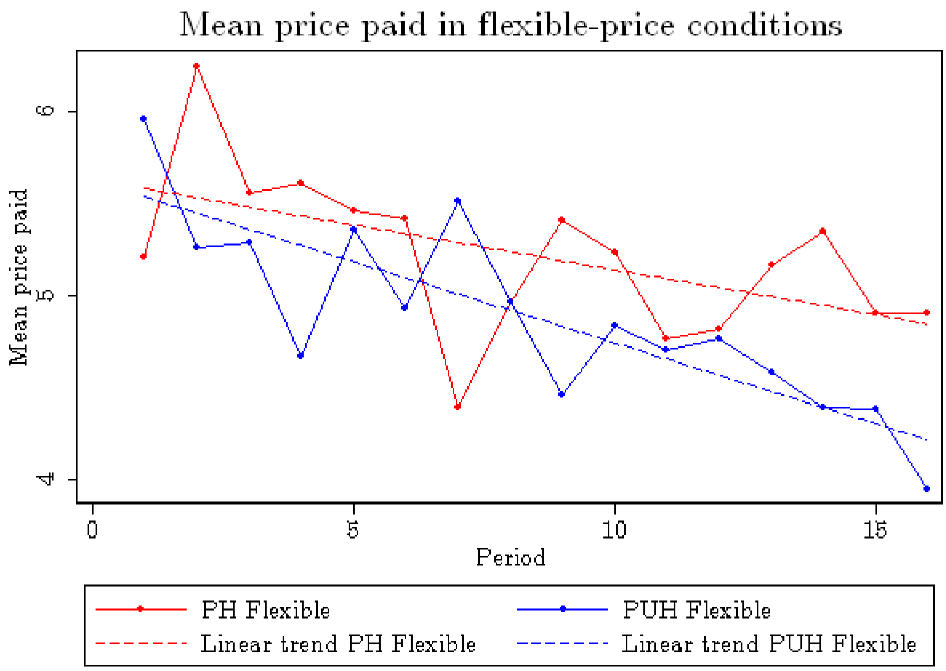


Figure 4.3: Development of prices paid over time in flexible-price conditions.

flexible. The descriptives even suggest that more customer information might lead to a higher level of undertreatment in periods 1–9. However, this difference is not statistically significant (see models (4) and (6) in *Table 4.3*; Mann Whitney U Test: $p = 0.916$). The higher undertreatment rate under public histories comes along with a lower average price paid by the customer in the public (5.15) compared to the private histories (5.38) condition. *Figure 4.3* illustrates the average price paid by customers over time and shows that for all except for two periods, the average price paid is lower under public than under private histories.

Price competition under public histories is more intense than under private histories as customers observe all customers' histories. Once an expert undertreated a customer in any of the previous periods, the expert has to balance his bad reputation not only towards the customer undertreated but towards all customers. Comparing our results to Dulleck et al. (2011) shows that private histories slightly reduce experts' undertreatment compared to a market where no reputation building is possible. In contrast, the public history mechanism induces a slightly higher level of

Table 4.5: Percentage of overcharging in periods 1–9.

		Reputation mechanism	
		Private histories	Public histories
Price system	Fixed	71.11%	41.24%
	Flexible	77.84%	86.54%

undertreatment. These results provide further evidence for the hypothesis that more customer information does not necessarily decrease the level of undertreatment (see *Table 4.2*).

Result 4.3 (Private vs. public histories under fixed prices: undertreatment)**.** *The level of undertreatment is non-significantly lower under public than under private histories if prices are fixed.*

In contrast to the flexible price treatments, the level of undertreatment is lower under public than under private histories if prices are fixed. This difference is not statistically significant (Mann Whitney U Test: $p = 0.103$). Given that the level of undertreatment is already low under private histories, it comes as no surprise that the additional customer information does not lead to a significant decrease in the rate of undertreatment. The additional customer information about whether other customers receive a sufficient treatment does, however, significantly increase the rate of interaction (Mann Whitney U Test: $p = 0.042$). While under private histories, customers interact in 73.26% of the cases, the rate of interaction amounts to 96.18% in the public histories condition.

4.6.2 Level of Overcharging

In the following, we present the results concerning the level of overcharging. *Table 4.5* provides an overview of the level of overcharging across conditions.

In order to disentangle the different drivers for the expert's incentive to overcharge and to control for individual heterogeneity, we again perform a random effects panel

Table 4.6: Random effects panel regressions on overcharging in periods 1–9.

	Panel probit				Panel OLS	
Overcharging	(1)	(2)	(3)	(4)	(5)	(6)
Period	0.043^{*}	0.042^{*}	0.041^{*}	0.046^{**}	0.011	0.012
	(0.022)	(0.022)	(0.022)	(0.022)	(0.008)	(0.008)
Private histories		0.301	0.240	-0.425^{**}	0.094	-0.086
		(0.191)	(0.162)	(0.213)	(0.073)	(0.060)
Fixed prices			-0.882^{***}	-1.492^{***}	-0.259^{***}	-0.440^{***}
			(0.163)	(0.213)	(0.078)	(0.116)
Private histories · fixed prices				1.324^{***}		0.388^{***}
				(0.305)		(0.137)
Intercept	0.486^{***}	0.340^{**}	0.759^{***}	1.045^{***}	0.728^{***}	0.808^{***}
	(0.146)	(0.170)	(0.176)	(0.186)	(0.068)	(0.063)
$R^2_{M\&Z}$	0.022	0.035	0.151	0.211	–	–
R^2	–	–	–	–	0.104	0.149
Observations	705	705	705	705	705	705

Standard errors are clustered on the individual level for panel probit regressions (Note: clustering for panel probit regressions on a different level than the individuals' is not implemented). Standard errors are robust and clustered on the market level for panel OLS regressions. Standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. p-values are based on two-tailed tests.

probit regression with standard errors clustered on the individual level and a random effects panel OLS regression with robust standard errors clustered on the market level. Besides the conditions, we control for the period in which an interaction takes place. We also add the interaction effect of the conditions. Thus, the specification is as follows where c_i denotes the random intercept of individual i and u_{it} denotes the idiosyncratic error term for individual i in period t:

$$\begin{aligned} overcharging_{it} = \beta_0 + \beta_1 period_{it} + \beta_2 private\_histories_{it} + \beta_3 fixed\_prices_{it} \\ + \beta_4 private\_histories_{it} \cdot fixed\_prices_{it} + c_i + u_{it}. \end{aligned} \tag{4.4}$$

Table 4.6 displays our regression results.

Result 4.4 (Flexible vs. fixed prices: overcharging). *The level of overcharging is significantly lower under fixed than under flexible prices.*

In line with *Hypothesis 4.4*, overcharging is significantly lower under fixed than under flexible prices (see models (3) and (5) in *Table 4.6*; Mann Whitney U Test: $p = 0.019$).[49] Following the OLS estimates, this difference in the probability of being overcharged amounts to 25.9 percentage points. If prices are flexible, competition drives down prices (see *Figure 4.2* and *4.3*). A possible explanation why experts overcharge more often if prices are flexible may be that experts try to compensate the losses due to the lower prices. Thus, price competition again undermines the reputation mechanism under flexible prices. The fact that customers return significantly more often to an expert who overcharged in the flexible- than in the fixed-price condition supports this argument (Mann-Whitney U Test: $p < 0.001$).[50]

Result 4.5 (Private vs. public histories under flexible prices: overcharging). *The level of overcharging is weak significantly higher under public than under private histories if prices are flexible.*

We find weak evidence that the level of overcharging is higher under public than under private histories if prices are flexible (Mann Whitney U Test: $p = 0.093$).[51] Price competition is stronger under public than under private histories because a customer observes the price each customer was charged in the previous periods and not only the price she was charged herself. Compensating the lower prices by higher frequencies of overcharging thus leads to an increased level of overcharging under public histories.

[49] Note that the difference in the level of overcharging between flexible and fixed prices is driven by the difference between the public history conditions (Mann Whitney U Test: $p = 0.006$). The difference between the private history conditions is not statistically significant (Mann Whitney U Test: $p = 0.834$).

[50] Note, however, that the customer does not directly observe the overcharging but only the relation of the prices charged. She can thereby form beliefs about an expert's overcharging behavior.

[51] Note that the panel probit regression supports this result on a 5% significance level whereas the panel OLS regression does not find a statistically significant difference.

Result 4.6 (Private vs. public histories under fixed prices: overcharging)**.** *The level of overcharging is weak significantly lower under public than under private histories if prices are fixed.*

This result on overcharging again follows the pattern we find for the level of undertreatment: in a fixed-price regime, there is less overcharging under public than under private histories. Looking at the descriptives in *Table 4.5*, this difference in the level of overcharging amounts to about 30 percentage points. According to the Mann Whitney U Test, however, the difference is only weakly significant ($p = 0.066$). The reason is that six out of the eight markets under private histories are characterized by high overcharging levels while the remaining two exhibit a relatively low level of overcharging.

The difference in the level of overcharging between public and private histories is due to the fact that customers can observe whether other experts charged the price for the minor or the major treatment in previous periods under public histories. Thus, experts have an incentive to keep up their reputation as an honestly charging expert in the first periods. Note that we virtually find no undercharging under public histories which is in strong contrast to the predicted expert behavior in the reputation equilibrium without overcharging (see *Lemma 4.5*).

4.7 Conclusion

This paper analyzes the impact of two reputation mechanisms under fixed prices and price competition on an expert's incentive to defraud. We make use of a 2×2 factorial design to test our predictions experimentally. The two main results of the paper are: firstly, under fixed prices, the level of undertreatment and the level of overcharging are considerably lower than under flexible prices. Under flexible prices, we observe a price pressure that undermines reputation building: experts who undertreated in previous periods try to balance their bad reputation by offering low prices. The lower prices in turn make it less attractive for the expert to treat sufficiently and to charge honestly. Secondly, we find weak evidence that the level of overcharging is higher under public than under private histories if prices are flexible while the opposite holds when prices are fixed. We find the same pattern for the level of undertreatment when comparing public with private histories but differences are not statistically significant. The additional customer information intensifies the price competition between experts because experts who defrauded have to balance their reputation not only towards the customer they defrauded but towards all customers. Hence, we observe more fraud under public than under private histories if prices are flexible.

Our results thus provide first indicative evidence that reputation building might increase the level of fraud in a market with asymmetric information. This result is in line with Ely and Välimäki (2003) and Ely et al. (2008). However, the mechanism at work differs: it is the expert's attempt to reach a history of treatments reflecting the ex-ante probabilities that may induce additional undertreatment in their set-up. In contrast, our results are driven by the intense price competition.[52]

The implications of our results are twofold as the introduction of public histories does not necessarily decrease the level of fraud but may even lead to an increase. In a market with fixed prices, feedback platforms such as the *Arztnavigator* seem to be an appropriate instrument to reduce the fraud level. In markets with flexible prices, however, more customer information may lead to adverse effects with respect to the fraud level.

In future research, it firstly remains to show how robust the increase in the level of fraud is when varying between private and public histories under flexible prices. Secondly, the robustness of our results against changes in the exogenously given price structure remains to be investigated. Dulleck et al. (2012) provide some evidence that the level of undertreatment is fairly robust to the price structure in a market without reputational concerns.

[52] Whereas we assume that customers cannot observe the treatment, Ely and Välimäki (2003) and Ely et al. (2008) assume the opposite. This enables the expert to build up a reputation based on his treatment history.

4.8 Appendix A: Proofs

In the following proofs, we assume that if customers are indifferent between visiting and not visiting an expert, the customer opts for the visit. Experts who are indifferent between undertreating and not undertreating do not undertreat.

For each class of equilibria (no-reputation and reputation equilibria) two types of equilibria may exist: an equilibrium where customers randomize with equal probability among experts and an equilibrium where customers coordinate on an expert in the first period. We outline both types of equilibria in the following if they exist.

4.8.1 Proof of Lemma 4.2

4.8.1.1 Reputation Equilibria in which Customers Randomize

We first comment on the reputation equilibria in which customers randomize between experts in the first period.

Public Histories The strategies and beliefs as well as the corresponding proof for the reputation equilibrium under public histories in which customers randomize are outlined in Dulleck et al. (2011). Note that the strategies and beliefs presented below partially build upon this equilibrium.

Private Histories Under private histories a reputation equilibrium where customers randomize between experts does not exist. This is because customers cannot coordinate on the expert with the strict majority of customers because customers do not observe how many customers an expert serves. Without customers observing which expert serves the most customers and without customer coordination, it is optimal for experts to undertreat customers unless they face four customers. If experts post prices as outlined in *Lemma 4.2* and an expert serves three customers, his maximum additional future payoff from treating customers sufficiently amounts to $7(3 - 1.6) = 9.8$ in period 9 (which is the last period where reputational concerns may play a role), while the maximum additional current payoff from deviating amounts to $3((5-2)-(5-6)) = 12$. Thus, experts always provide the minor treatment if they serve three or less customers. The expected payoff from interacting $(0.5+)5 - 0.5(1-1/64) \cdot 5 = 0.078125$ is lower than the outside option of 1.6 (where $1/64$ is the probability that all customers choose the same expert in the first period).

4.8.1.2 Reputation Equilibrium in which Customers Coordinate

In the following, we assume that customers coordinate on one of the experts in the first period. We refer to the four experts as expert A, B, C, and D.

Public Histories The strategies and beliefs in the reputation equilibrium with customers coordinating under public histories are as follows.

Customers' beliefs: Each customer believes that experts always charge p_H. Each customer believes to be treated sufficiently if and only if (i) she is treated under a price menu $\{n.d., 5\}$ and the expert has at least two customers, (ii) the expert has only undertreated a customer in situation where all experts served exactly one customer, and (iii) the game is in periods 1–9. Otherwise, each customer believes to get the minor treatment.

Customers' strategy: Each customer visits among the experts posting a price menu $\{n.d., 5\}$ the same expert as all other customers (in the following expert A) in the first period. In periods 2–9, if expert A did not undertreat any of the customers in the previous period and expert A posts a price menu a $\{n.d., 5\}$ in the current period, customers return to expert A. If expert A undertreated any of the customers in the previous period or posts a price menu different from $\{n.d., 5\}$ in the current period, all customers coordinate to another expert (in the following expert B) among the experts posting a price menu $\{n.d., 5\}$. If there is no expert posting $\{n.d., 5\}$, customers randomize between those experts posting $\{n.d., 3\}$. If there is no expert posting $\{n.d., 3\}$, customers do not interact. Customers' strategy when visiting experts B, C and D is according to the above strategy at expert A. If there is no expert who never undertreated, customers do not interact. In periods 10–16, customers choose expert A if he never undertreated in periods 1–9. If expert A

undertreated in any period 1–9, customers visit expert B if he never undertreated in any period 1–9; and so forth. If there is no expert who never undertreated, customers randomize between experts with equal probability in periods 10–16.

Expert Strategy: In the first nine periods, all experts post $\{n.d., 5\}$. Each expert serves his customers sufficiently if he has two or more customers and provides the minor treatment otherwise. In periods 10–16, all experts post the price menu $\{n.d., 3\}$ and always provide the minor treatment if one seller had strictly the most customers in period 9. Otherwise all experts play the strategy described *Lemma 4.1*.

Verification:
We now verify that the given strategies and beliefs form a perfect Bayesian equilibrium. We first show that customers' strategies are rational. In periods 10–16, if customers interact, they receive an expected payoff of $0.5(10-3)+0.5(0-3)=2$ which is larger than their outside option of 1.6. In periods 1–9, given expert behavior, it is optimal to interact as the expected payoff of $10-5=5$ is larger than the one from not interacting (1.6). If customers return to play the above described randomization equilibrium, customers' strategy off-equilibrium is optimal as shown in the public histories reputation equilibrium with randomization.

In the following, we show that experts' strategy is rational. In periods 10–16 it is optimal to always provide the minor treatment because the maximum additional future payoff of treating sufficiently $6 \cdot 2.4 = 14.4$ (in period 10) is lower than the maximum current payoff from deviating $4((5-2)-(5-6)) = 16$. In period 9, if the expert serves four customers, the expert's maximum additional future payoff of treating customers sufficiently amounts to $7 \cdot 2.4 = 16.8$ whereas the maximum additional current payoff from deviating is $4((5-2)-(5-6)) = 16$. If the expert serves three customers, his maximum additional future payoff of treating customers sufficiently also amounts to $7(4-1.6) = 16.8$ (because the single customer will visit this expert in periods 10–16) whereas the maximum additional current profit from deviating is $3((5-2)-(5-6)) = 12$. If an expert serves two customers in period 9, the maximum current payoff from deviating amounts to $2((5-2)-(5-$

6)) = 8. Whether the expert will serve all four customers in periods 10–16 depends on whether the expert serves the strict majority of customers in period 9. Given customers' strategies, an expert should never serve only two customers. Hence, Bayes' rule cannot be applied to calculate the probability with which an expert expects to be the expert with the strict majority of customers. We assume that the expert believes to serve the strict majority of customers with probability 1. Given these beliefs, the expert's maximum additional future payoff of treating sufficiently amounts to 16.8 because all customers will choose this expert in periods 10–16. In periods 1–8, future payoff of treating customers sufficiently is larger so that deviation incentives are lower. Hence, the expert will treat customers sufficiently if he serves at least two customers under the above beliefs. In case the expert only serves one customer, the maximum additional future payoff from treating sufficiently amounts to $7(1 - 1.6) = -2.8$ while the maximum current payoff from deviating is 4. Thus, the expert always provides the minor treatment if he serves a single customer. Given that all other experts charge the price p_H for both treatments in all 16 periods, it is optimal for the individual expert to also charge p_H. Consequently, experts' strategies are rational.

Private Histories The strategies and beliefs in the reputation equilibrium with customers coordinating under private histories are as follows.

Customers' beliefs: Each customer expects to be charged p_H in any of the periods. Each customer believes to be treated sufficiently if and only if (i) she is treated under a price menu $\{n.d., 5\}$ and the expert has four customers, (ii) if the expert has never undertreated the customer before, and (iii) the game is in periods 1–9. Otherwise, each customer believes to get the minor treatment.

Customers' strategy: Each customer visits among the experts that post a price menu $\{n.d, 5\}$ the same expert as all other customers (in the following expert A) in the first period. In periods 2–9, if expert A did not undertreat the customer in the previous period and posts $\{n.d, 5\}$, the customer returns to expert A. If expert A undertreated the customer in the previous period or posts a price vector different from $\{n.d, 5\}$, the customer refrains from interacting. In periods 10–16, if expert A did not undertreat the customer in any period 2–9, the customer stays with expert A. If expert A undertreated the customer in any period 2–9, the customer randomizes between the remaining three experts with equal probability in periods 10–16.

Experts' strategy: Experts post price vectors $\{n.d., 5\}$ in periods 1–9 and $\{n.d., 3\}$ in periods 10–16. Each expert treats his customers sufficiently in periods 1–9 if he serves all four customers and provides the minor treatment otherwise. In periods 10–16, experts always provide the minor treatment. Experts always charge p_H.

Verification:

We now verify that the given strategies and beliefs form a perfect Bayesian equilibrium. We first show that customers' strategies are rational. In periods 10–16, if customers interact, they receive an expected payoff of $0.5(10-3)+0.5(0-3) = 2$ which is larger than their outside option of 1.6. In periods 1–9, given expert behavior, i.e., always sufficient treatment, it is optimal to interact as the expected payoff of $10-5 = 5$ is larger than the one from not interacting (1.6). If expert A undertreats, it is optimal for the customer to refrain from interacting because the outside option of 1.6 is larger than the expected payoff from interacting $(0.5(10-5)+0.5(0-5) = 0)$.

In the following, we show that experts' strategies are rational. In periods 10–16 it is optimal to always provide the minor treatment because the maximum additional future payoff of treating sufficiently $6 \cdot 2.4 = 14.4$ (in period 10) is lower than the maximum current payoff from deviating $4((5-2)-(5-6)) = 16$. In period 9, if the expert serves four customers, the expert's maximum additional future payoff of treating customers sufficiently amounts to $7 \cdot 2.4 = 16.8$ whereas the maximum additional current payoff from deviating is $4((5-2)-(5-6)) = 16$. In periods 1–8, future payoff of treating customers sufficiently is larger so that deviation incentives are lower. If the expert serves three or less customers, his maximum additional future payoff of treating customers sufficiently amounts to $7(3-1.6) = 9.8$ (in case he serves three customers) which is less than the maximum additional current profit from deviating $3((5-2)-(5-6)) = 12$.[53] Thus, it is optimal for the expert to provide the minor treatment. Given that all other experts charge the price p_H for both treatments in all 16 periods, it is optimal for the individual expert to also charge p_H. Hence, experts' strategy is rational.

4.8.2 Proof of Lemma 4.3

Both types of equilibria (with customer randomization and customer coordination) described in *Lemma 4.3* are characterized as follows:

Customers' beliefs: Each customer believes to always receive the minor treatment and to always be charged p_H.

Customers' strategy: Customers do not interact in periods 1–9. Customers randomize between experts in each period respectively coordinate in any (arbitrary) way on the experts in periods 10–16.

Experts' strategy: Experts always provide the minor treatment and always charge p_H.

Verification:

We now verify that the above outlined strategies and beliefs form a perfect Bayesian equilibrium. Customers' behavior is rational because their expected payoff from interaction in periods 1–9 amounts to $0.5(10-8)+0.5(0-8) = -3$ which is less than the outside option of 1.6. In periods 10–16, if customers interact, they receive an expected payoff of $0.5(10-3)+0.5(0-3) = 2$ which is larger than their outside option of 1.6. Given the customers' behavior, experts' strategies are optimal because their payoff from always providing the minor treatment at the price p_H is larger than treating sufficiently. Note that the expert cannot decide not to participate.

[53]Note that the fourth customer in the market will be undertreated because she ends up as a single customer at an expert in period 9. As the fourth customer cannot observe whether the expert serving three customers undertreated his customers, her behavior in periods 10–16 is not changed by the treatment decision of the expert serving three customers in period 9. Hence, the fourth customer is irrelevant when determining whether the expert serving three customers deviates or sticks to the equilibrium strategy.

4.8.3 Proof of Lemma 4.4

The players' strategies and beliefs in the reputation equilibria without overcharging in a market with fixed prices are outlined below. The reputation equilibria require that customers coordinate on one expert in the first period. Given the strategies and beliefs of reputation equilibria under flexible prices in which customers randomize, customers would face a lower payoff than the outside option under fixed prices. Thus, if interaction is still observed, customers must coordinate on one of the experts. We refer to the four experts as expert A, B, C, and D.

Public Histories The strategies and beliefs in the reputation equilibrium with customer coordination under public histories are as follows.

Customers' beliefs: Each customer expects to be charged p_H in any of the periods. Each customer believes to be treated sufficiently if and only if (i) the expert serves at least two customers, (ii) the expert only undertreated a customer in situation where all experts served one customer, and (iii) the game is in periods 1–9. Otherwise, each customer believe to receive the minor treatment.

Customers' strategy: Each customer visits the same expert (in the following expert A) as all other customers in the first period. In periods 2–9, if expert A did not undertreat any of the customers in the previous period, customers return to expert A. If expert A undertreated in the previous period, all customers visit expert B. If expert B did not undertreat any of the customers in the previous period, customers return to expert B whereas if he undertreated in the previous period, customers choose expert C; and so forth. If there is no expert who never undertreated, customers do not interact. In periods 10–16, customers choose expert A if he never undertreated. If expert A undertreated in any period 1–9, customers visit expert B if he never undertreated; and so forth. If there is no expert who never undertreated, customers randomize between experts with equal probability in periods 10–16.

Experts' strategy: Each expert treats his customers sufficiently in periods 1–9 if he serves two or more customers and provides the minor treatment otherwise. In periods 10–16, experts always provide the minor treatment. Experts always charge p_H.

Verification:
We now verify that the above described strategies and beliefs form a perfect Bayesian equilibrium. We first show that customers' strategies are rational. In periods 10–16, if customers interact, they receive an expected payoff of $0.5(10-3)+0.5(0-3) = 2$ which is larger than their outside option of 1.6. In periods 1–9, given expert behavior, i.e., always sufficient treatment, it is optimal to interact as the expected payoff of $10-8 = 2$ is larger than the one from not interacting (1.6).

In the following, we show that experts' strategy is rational. In periods 10–16 it is optimal to always provide the minor treatment because the maximum additional future payoff of treating sufficiently $6 \cdot 2.4 = 14.4$ (in period 10) is lower than the maximum current payoff from deviating $4((8-2)-(8-6)) = 16$. In period 9, if an expert serves four customers, the expert's maximum additional future payoff of treating customers sufficiently amounts to $7 \cdot 2.4 = 16.8$ whereas the maximum additional current payoff from deviating is $4((8-2)-(8-6)) = 16$. If the expert serves three customers, his maximum additional future payoff of treating customers sufficiently also amounts to $7(4-1.6) = 16.8$ (because the single customer will visit this expert in periods 10–16) whereas the maximum additional current profit from deviating is $3((8-2)-(8-6)) = 12$. If an expert serves two customers in period 9, the maximum current payoff from deviating amounts to $2((8-2)-(8-6)) = 8$. Whether the expert serving two customers in period 9 will serve all four customers in periods 10–16 depends on whether the expert serves the strict majority of customers in period 9. Given customers' strategies, an expert should never serve only two customers. Hence, Bayes' rule cannot be applied to calculate the probability with which an expert expects to be the expert with the strict majority of customers. We assume that the expert believes to serve the strict majority of customers with probability 1. Given these beliefs, the expert's maximum additional future payoff of treating sufficiently amounts to 16.8 because all customers will choose this expert in periods 10–16. In periods 1–8, future payoff of treating customers sufficiently is larger so that deviation incentives are lower. Hence, the expert will treat customers sufficiently if he serves at least two customers under the above beliefs. In case the expert only serves one customer, the maximum additional future payoff from treating sufficiently amounts to $7(1 - 1.6) = -2.8$ while the maximum current payoff from deviating is 4. Thus, the expert always provides the minor treatment if he serves a single customer. Given that all other experts charge the price p_H for both treatments in all 16 periods, it is optimal for the individual expert to also charge p_H. Consequently, experts' strategies are rational.

Private Histories The strategies and beliefs in the reputation equilibrium with customers coordinating under private histories are as follows.

Customers' beliefs: Each customer expects to be charged p_H in any of the periods. Each customer believes to be treated sufficiently if and only if (i) the expert has four customers, (ii) the expert has never undertreated the customer before, and if (iii) the game is in periods 1–9. Otherwise, each customer expects to receive a minor treatment.

Customers' strategy: Each customer visits the same expert as all other customers (in the following expert A) in the first period. In periods 2–9, if expert A did not undertreat any of the customers in the previous period, customers return to expert A. If expert A undertreated in the previous period, customers refrain from interacting. In periods 10–16, if expert A did not undertreat the customer in any period 2–9, customers stay with expert A. If expert A undertreated the customer in any period 2–9, customers randomize between experts with equal probability in periods 10–16.

Experts' strategy: Each expert treats his customers sufficiently in periods 1–9 if he serves all four customers and provides the minor treatment otherwise. In periods 10–16, experts always provide a minor treatment. Experts always charge p_H.

Verification:
We now verify that the above described strategies and beliefs form a perfect Bayesian equilibrium. We first show that customers' strategies are rational. In periods 10–16, if customers interact, they receive an expected payoff of $0.5(10-3)+0.5(0-3) = 2$ which is larger than their outside option of 1.6. In periods 1–9, given expert behavior, it is optimal to interact as the expected payoff of $10 - 8 = 2$ is larger than the one from not interacting (1.6).

In the following, we show that experts' strategy is rational. In periods 10–16 it is optimal to always provide the minor treatment because the maximum additional future payoff of treating sufficiently $6 \cdot 2.4 = 14.4$ (in period 10) is lower than the maximum current payoff from deviating $4((8-2)-(8-6)) = 16$. In period 9, if the expert serves four customers, the expert's maximum additional future payoff of treating customers sufficiently amounts to $7 \cdot 2.4 = 16.8$ whereas the maximum additional current payoff from deviating is $4((8-2)-(8-6)) = 16$. In periods 1–8, future payoff of treating customers sufficiently is larger so that deviation incentives are lower. If the experts serves three or less customers, his maximum additional future payoff of treating customers sufficiently amounts to 9.8 (in case he serves three customers) which is less than the maximum additional current profit from deviating $3((8-2)-(8-6)) = 12$.[54] Thus, it is optimal for the expert to provide the minor treatment. Given that all other experts charge the price p_H for both treatments in all 16 periods, it is optimal for the individual expert to also charge p_H. Hence, experts' strategy is rational.

4.8.4 Proof of Lemma 4.5

The strategies and beliefs of the players in the reputation equilibrium without overcharging in a market with fixed prices are outlined below. We refer to the four experts as expert A, B, C, and D.

[54] Note that the single customer is again irrelevant for the above analysis.

Customers' beliefs: Each customer believes to be charged p_L if (i) the expert serves at least two customers, (ii) the expert has never charged p_H in any of the previous periods and (iii) the game is in period 1–7. Otherwise, each customer expects to be charged p_H. Each customer believes to be treated sufficiently if and only if (i) the expert serves at least two customers, (ii) the expert has never undertreated in any of the previous periods, and (iii) the game is in periods 1–9. Otherwise, each customer expects to receive the minor treatment.

Customers' strategy: Each customer visits the same expert as all other customers (in the following expert A) in the first period. If expert A did not undertreat any of the customers and did not charge p_H in the previous period, customers return to expert A. If expert A undertreated in the previous period or charged p_H, all customers visit expert B. If expert B did not undertreat any of the customers and did not charge p_H in the previous period, customers return to expert B; otherwise, all customers visit expert C. If expert C did not undertreat any of the customers and did not charge p_H in the previous period, customers return to expert C; otherwise, all customers visit expert D. If expert D did not undertreat any of the customers and did not charge p_H in the previous period, customers return to expert D; otherwise, all customers visit expert A if expert A has not undertreated in any of previous periods; otherwise, all customers visit expert B; and so forth. If there is no expert who has not undertreated, customers do not interact. In periods 10–16, customers choose expert A if he has not undertreated customers in periods 1–9 and has only charged p_L in periods 1–7. Otherwise, customers choose expert B if he has not undertreated customers in periods 1–9 and has only charged p_L in periods 1–7; and so forth. If there is no expert who never undertreated, customers randomize between experts with equal probability in periods 10–16.

Experts' strategy: Each expert treats his customers sufficiently in periods 1–9 if he serves two or more customers and provides the minor treatment otherwise. Experts always provide the minor treatment in periods 10–16. In periods 1–7, each expert charges each customer p_L if he serves at least two customers and charges p_H otherwise. In periods 8–16, experts always charge p_H.

Verification:

We now verify that the above described strategies and beliefs form a perfect Bayesian equilibrium. We first show that customers' strategies are rational. Customers interact in periods 1–8 because their expected payoff from interacting amounts to $10 - 4 = 6$ which is more than the outside option of 1.6. Customers interact in period 9 because $0.5(10-8)+0.5(10-8) = 2 > 1.6$. Customers interact in periods 10–16 because $0.5(10-3)+0.5(0-3) = 2 > 1.6$.

In the following, we show that experts' strategy is rational. In periods 10–16 it is optimal to always provide the minor treatment because the maximum additional future payoff of treating sufficiently $6 \cdot 2.4 = 14.4$ (in period 10) is lower than the maximum current payoff from deviating $4((8-2)-(8-6)) = 16$. In period 9, if the expert serves four customers, the expert's maximum additional future payoff of treating customers sufficiently amounts to $7 \cdot 2.4 = 16.8$ whereas the maximum additional current payoff from deviating is $4((8-2)-(8-6)) = 16$. If the expert serves three customers, his maximum additional future payoff of treating customers sufficiently also amounts to $7(4-1.6) = 16.8$ (because the single customer will visit this expert in periods 10–16) whereas the maximum additional current profit from deviating is $3((8-2)-(8-6)) = 12$. If an expert serves two customers in period 9, the maximum current payoff from deviating amounts to $2((8-2)-(8-6)) = 8$. Whether the expert will serve all four customers in periods 10–16 depends on whether the expert serves the strict majority of customers in period 9. Given customers' strategies, an expert should never serve only two customers. Hence, Bayes' rule cannot be applied to calculate the probability with which an expert expects to be the expert with the strict majority of customers. We assume that the expert believes to serve the strict majority of customers with probability 1. Given these beliefs, the expert's maximum additional future payoff of treating sufficiently amounts to 16.8 because all customers will choose this expert in periods 10–16. In periods 8 and 9, future payoff of treating customers sufficiently is larger so that deviation incentives are lower. Hence, the expert will treat customers sufficiently if he serves at least two customers under the above beliefs. In case the expert only serves one customer, the maximum additional future payoff from treating sufficiently amounts to $7(1 - 1.6) = -2.8$ while the maximum current payoff from deviating is 4.

Next, we consider check experts' incentive to deviate from their strategy in periods 1–7. Note that experts' incentive to deviate is largest in period 1 and not in period 7. This is because experts make zero profits if they play according to equilibrium strategy in periods 1–7 while a deviation leads to profits of 1.6 (outside option). In period 1, experts' behavior is rational because the maximum additional current payoff from deviating—if all four customers have a major problem, the expert charges p_H but provides the minor treatment—amounts to $4((8-2)-(4-6)) = 32$ while the maximum expected future payoff from charging p_L and treating sufficiently amounts to $2 \cdot 4(0.5(8-6) + 0.5(8-2)) - 9 \cdot 1.6 + 7 \cdot 2.4 = 34.4$.[55] For the case of three and two customers, the same reasoning applies as for the deviation in period 9. Hence, experts charge p_L if they face at least two customers given the above outlined belief. In periods 2–7, the incentive to deviate is lower than in period 1. Thus, the experts' behavior is rational.

[55]In periods 1–9, the expert sticking to the equilibrium strategy gives up the outside option of 1.6. In periods 8 and 9, the expert can charge all four customers the major treatment although in expectation only two customers need the major treatment. In periods 10–16, the expert's additional expected future profit amounts to $7(4 - 1.6)$.

4.9 Appendix B: Screenshots of Feedback Systems

4.9.1 Feedback System in a Fixed Price Market

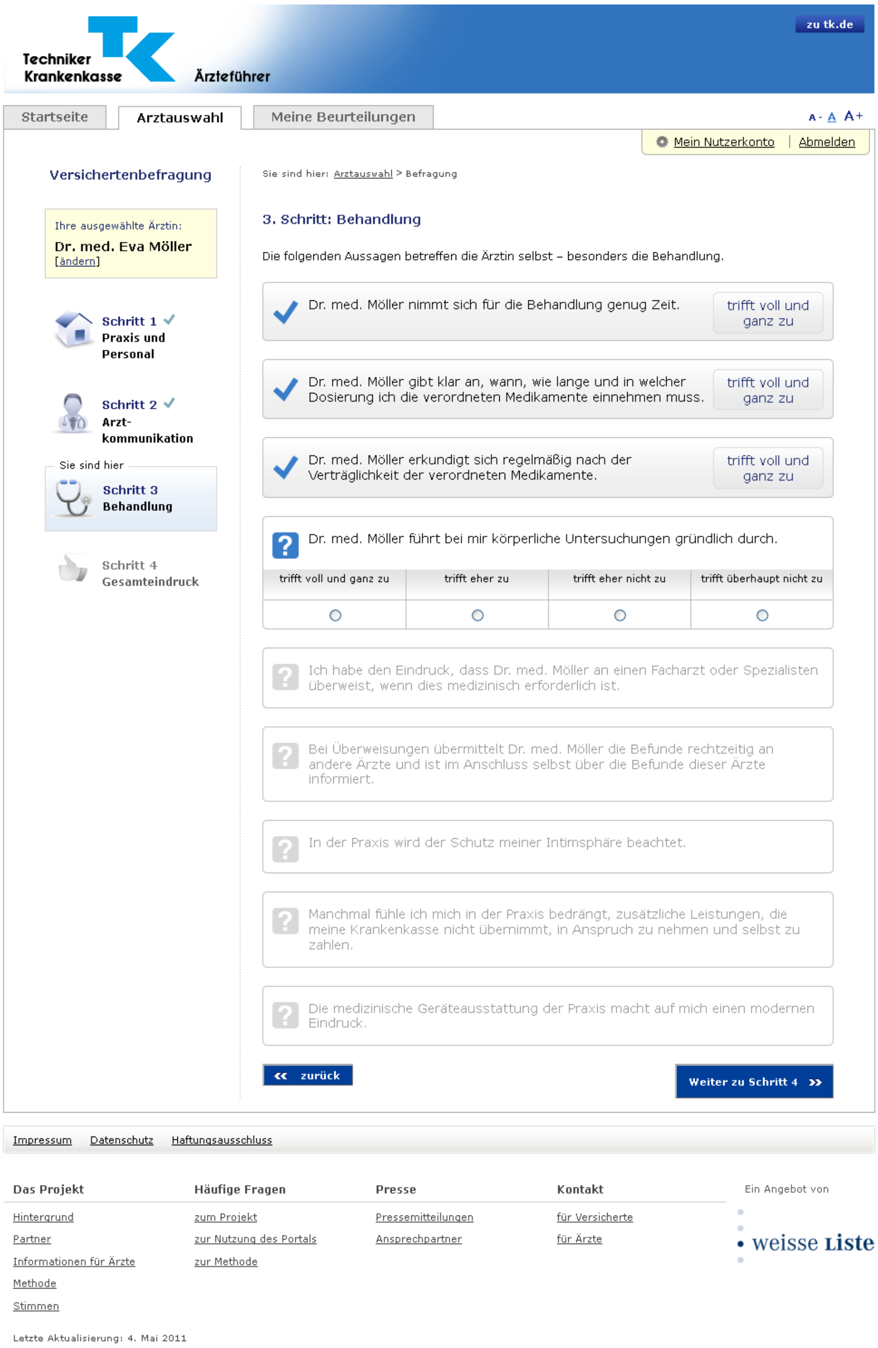


Figure 4.4: Patient feedback at the *Arztnavigator*. Source: `https://weisse-liste.arzt-versichertenbefragung.tk.de/`, accessed on July 18, 2012.

4.9.2 Feedback System in a Flexible Price Market

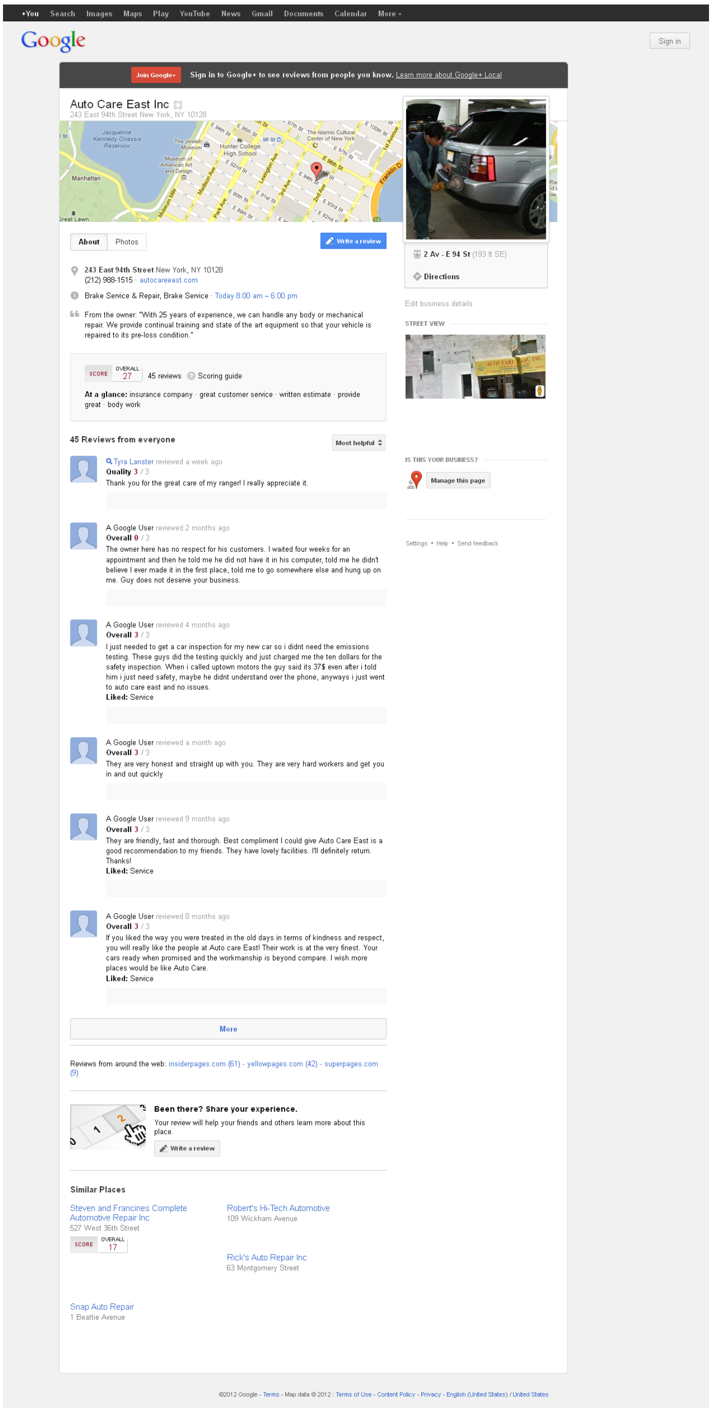


Figure 4.5: Car repair shop rating at *Google Maps*.
Source: `https://plus.google.com/109459300714062123468/about?gl=US&hl=en`, accessed on July 18, 2012.

4.10 Appendix C: Instructions

In the following, we present the instructions for the public histories under flexible prices treatment. The instructions are taken from Dulleck et al. (2011) and have been adapted for our purposes.

ANLEITUNG ZUM EXPERIMENT

Herzlichen Dank für Ihre Teilnahme am Experiment. Bitte sprechen Sie bis zum Ende des Experiments nicht mehr mit anderen Teilnehmern.

2 Rollen und 16 Runden
Dieses Experiment besteht aus **16 Runden**, die jeweils die gleiche Abfolge an Entscheidungen haben. Die Abfolge der Entscheidungen wird unten ausführlich erklärt.
Es gibt im Experiment 2 Rollen: **Spieler A** und **Spieler B**. Zu Beginn des Experiments bekommen Sie eine dieser Rollen zufällig zugelost und behalten diese Rolle für das gesamte Experiment. Auf dem ersten Bildschirm des Experiments sehen Sie, welche Rolle Sie haben. Diese Rolle bleibt für alle Spielrunden gleich.
In Ihrer Gruppe sind 4 Spieler A und 4 Spieler B. Die Spieler jeder Rolle bekommen eine Nummer. Sind Sie ein Spieler B, dann sind Ihre potentiellen Interaktionspartner die Spieler A1, A2, A3 und A4. Sind Sie hingegen ein Spieler A, dann sind Ihre potentiellen Interaktionspartner die Spieler B1, B2, B3 und B4.
Die Nummern der **Spieler** sind **fix**. Das heißt, dass zum Beispiel hinter der Nummer „A1" oder hinter der Nummer „B3" immer dieselbe Person steht. Spieler A erfährt zu keinem Zeitpunkt, mit welchem/welchen Spieler/n B (B1-B4) er interagiert.

Alle Experimentteilnehmer erhalten die gleichen Informationen bezüglich der Regeln des Spiels, inklusive der Kosten und Auszahlungen an beide Spieler.

Überblick über die Entscheidungen in einer Runde
Jede einzelne Runde besteht aus maximal 4 Entscheidungen, die hintereinander getroffen werden. Die Entscheidungen 1, 3 und 4 werden von Spieler A getroffen; die Entscheidung 2 wird von Spieler B getroffen.

Ablauf der Entscheidungen einer Runde (kurz gefasst)
1. Die Spieler A wählen Preise für die Aktionen 1 und 2.
2. Jeder Spieler B erfährt die von den 4 Spielern A (A1 bis A4) gewählten Preise. Dann entscheidet Spieler B, ob er mit einem Spieler A interagieren möchte. Es ist nur möglich, mit *einem* Spieler A zu interagieren. Falls Spieler B mit keinem Spieler A interagiert, endet diese Runde für ihn.

Falls Spieler B mit einem Spieler A interagiert ...

3. Der jeweilige Spieler A erhält die Information, ob einer oder mehrere Spieler B mit ihm interagieren. Es können maximal alle 4 Spieler B mit einem bestimmten Spieler A interagieren. Spieler A erfährt dann, welche Eigenschaften die Spieler B haben, die mit ihm interagieren. Es gibt zwei mögliche Eigenschaften: Eigenschaft 1 oder Eigenschaft 2. Diese Eigenschaft muss nicht identisch sein für die betreffenden Spieler B. Spieler A muss für jeden Spieler B, mit dem er interagiert, eine Aktion wählen: entweder Aktion 1 oder Aktion 2.
4. Spieler A verlangt von Spieler B den in Entscheidung 1 festgelegten Preis für eine der beiden Aktionen. Dabei muss der verlangte Preis nicht gleich dem Preis der in Entscheidung 3 gewählten Aktion sein, sondern es kann auch der Preis der anderen Aktion sein. Außerdem kann Spieler A von verschiedenen Spielern B unterschiedliche Preise verlangen.

Detaillierte Darstellung der Entscheidungen und ihrer Konsequenzen hinsichtlich der Auszahlungen

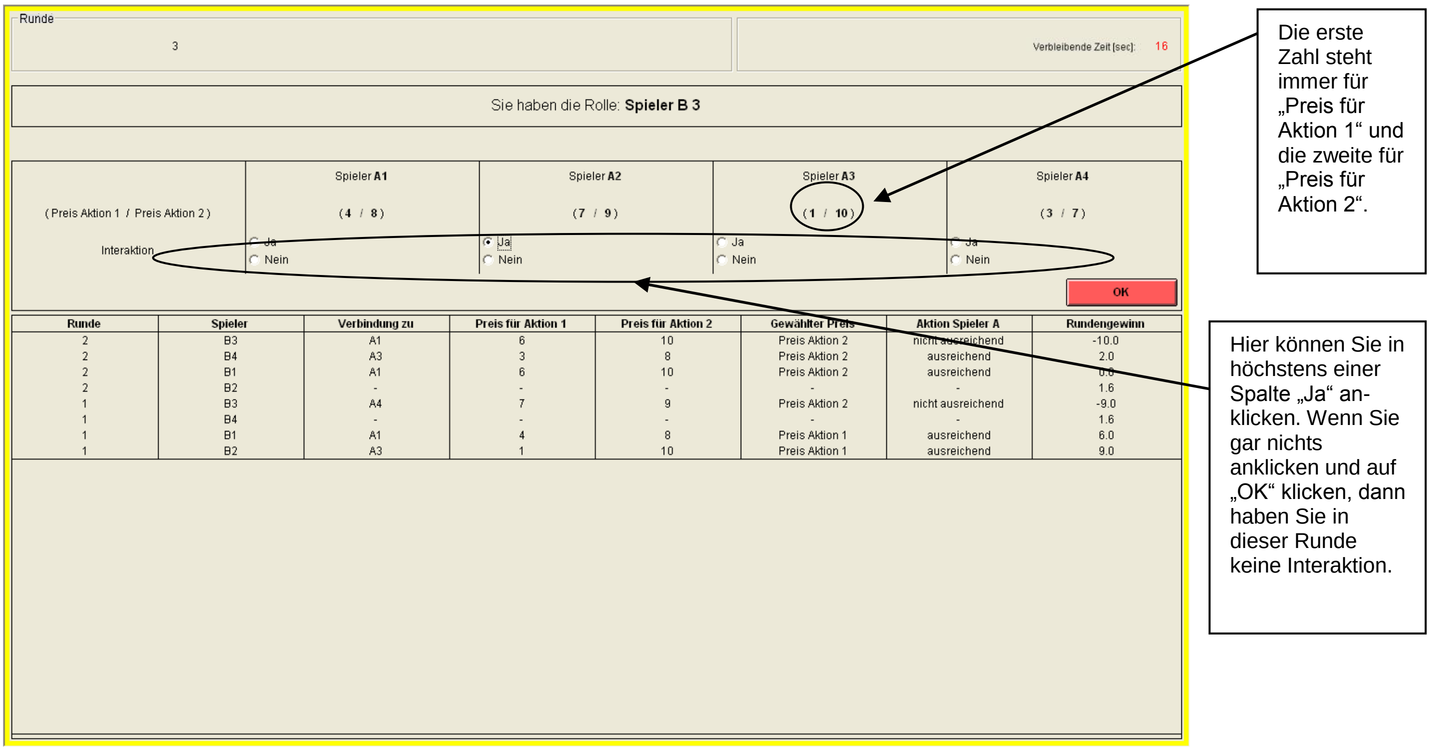
Entscheidung 1
Jeder Spieler A hat in Entscheidung 3 für den Fall einer Interaktion zwischen zwei Aktionen zu wählen, einer Aktion 1 und einer Aktion 2. Jede gewählte Aktion verursacht Kosten, die folgendermaßen fixiert sind:
Die **Aktion 1** verursacht **Kosten von 2 Punkten** (= experimentelle Währungseinheit) für Spieler A.
Die **Aktion 2** verursacht **Kosten von 6 Punkten** für Spieler A.
Für diese Aktionen kann Spieler A von jenen Spielern B, die mit ihm interagieren wollen, Preise verlangen. In **Entscheidung 1** muss jeder Spieler A diese **Preise für beide Aktionen festlegen**. Nur

(strikt) positive Preise in vollen Punkten von 1 Punkt bis maximal 11 Punkte sind möglich. D.h. die zulässigen Preise sind 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 oder 11.
Beachten Sie, dass der Preis für die Aktion 1 den Preis für die Aktion 2 nicht übersteigen darf.

Entscheidung 2
Spieler B erfährt die von allen 4 Spielern A in Entscheidung 1 gesetzten Preise. Dann entscheidet Spieler B, ob er mit einem der Spieler A interagieren möchte, und wenn ja, mit welchem.
Falls ja, dann bedeutet das, dass der entsprechende Spieler A in den Entscheidungen 3 und 4 eine Aktion wählen und dafür einen Preis verlangen kann (siehe unten). Spieler B wird aber **nicht** beobachten können, welche Aktion Spieler A wählt.
Falls nein, dann **endet** diese Runde für diesen Spieler B und er erhält als **Auszahlung für diese Runde 1,6 Punkte**.
Falls **keiner der Spieler B** mit einem bestimmten **Spieler A** interagieren möchte, erhält auch der betreffende Spieler A als **Auszahlung für diese Runde 1,6 Punkte**.

Auf der Folgeseite sehen Sie einen **exemplarischen Bildschirm** für die Entscheidung 2. Wenn Sie eine Interaktion mit einem bestimmten A-Spieler wünschen, dann klicken Sie bitte in der entsprechenden Spalte auf „Ja“ und bestätigen die Eingabe mit „OK“ (Sie müssen bei den anderen 3 A-Spielern dann nicht auf „Nein“ klicken). Wenn Sie überhaupt keine Interaktion wollen, dann müssen Sie nicht 4 Mal auf „Nein“ klicken, sondern können einfach OK bestätigen. (siehe Bildschirmerklärung).
In der unteren Hälfte des Bildschirms sehen Sie alle bisherigen Runden (aktuell ist Runde 3). Die Spalten bedeuten Folgendes:

- Runde: In welcher Runde etwas passiert ist
- Spieler: Um welchen Spieler B es sich handelt
- Verbindung zu: Hier sehen Sie, mit welchem Spieler A der jeweilige Spieler B interagiert hat (z.B. B4 in Runde 2 mit A3; „-“ falls keine Interaktion stattgefunden hat).
- Preis für Aktion 1: welchen Preis der jeweilige Spieler A für Aktion 1 festgesetzt hat (falls Sie eine Interaktion hatten; sonst steht „-“ wie z.B. bei B4 in Runde 1).
- Preis für Aktion 2: welchen Preis der jeweilige Spieler A für Aktion 2 festgesetzt hat.
- Gewählter Preis: „Preis Aktion 1“ bedeutet, dass der Preis für die Aktion 1 gewählt wurde (z.B. in Runde 1 von A1). „Preis Aktion 2“ bedeutet, dass der Preis für Aktion 2 gewählt wurde (z.B. in Runde 2 von A3). „-“ wird angezeigt bei keiner Interaktion.
- Aktion Spieler A: „ausreichend“ oder „nicht ausreichend“ (falls Interaktion stattgefunden hat) bzw. „-“ (falls keine Interaktion stattgefunden hat – wie in Runde 2 bei Spieler B2). (zur Erklärung siehe unten)
- Rundengewinn: Ihr Gewinn in Punkten in der betreffenden Runde. (zur Berechnung siehe unten)



Entscheidung 3

Vor der Entscheidung 3 (falls Spieler B in Entscheidung 2 „Ja“ gewählt hat) wird dem Spieler B zufällig eine Eigenschaft zugelost. **Spieler B** kann 2 Eigenschaften haben: **Eigenschaft 1** oder **Eigenschaft 2**. Die Eigenschaft wird **jede Runde neu** und auch für jeden Spieler B **unabhängig** zufällig bestimmt. Jeder Spieler B hat mit einer Wahrscheinlichkeit von **50% die Eigenschaft 1** und mit einer Wahrscheinlichkeit von **50% die Eigenschaft 2**. Stellen Sie sich in jeder Runde für jeden Spieler B einen Münzwurf vor. Wenn beispielsweise „Kopf“ kommt, dann hätte der entsprechende Spieler B die Eigenschaft 1, falls „Zahl“ kommt, hätte er die Eigenschaft 2.

Jeder Spieler A erfährt *vor* seiner Entscheidung 3 die **Eigenschaften aller jener Spieler B**, die mit diesem Spieler A interagieren wollen. Dann wählt Spieler A eine Aktion für jeden Spieler B, entweder Aktion 1 oder Aktion 2. Dabei kann die Aktion bei mehreren Spieler B auch unterschiedlich sein.
Eine **Aktion** ist unter folgenden Bedingungen für einen bestimmten Spieler B **ausreichend**:

a) Spieler B hat die Eigenschaft 1 und Spieler A wählt entweder die Aktion 1 oder die Aktion 2.
b) Spieler B hat die Eigenschaft 2 und Spieler A wählt die Aktion 2.

Eine Aktion ist **nicht ausreichend**, wenn Spieler B die Eigenschaft 2 hat, aber Spieler A die Aktion 1 wählt.

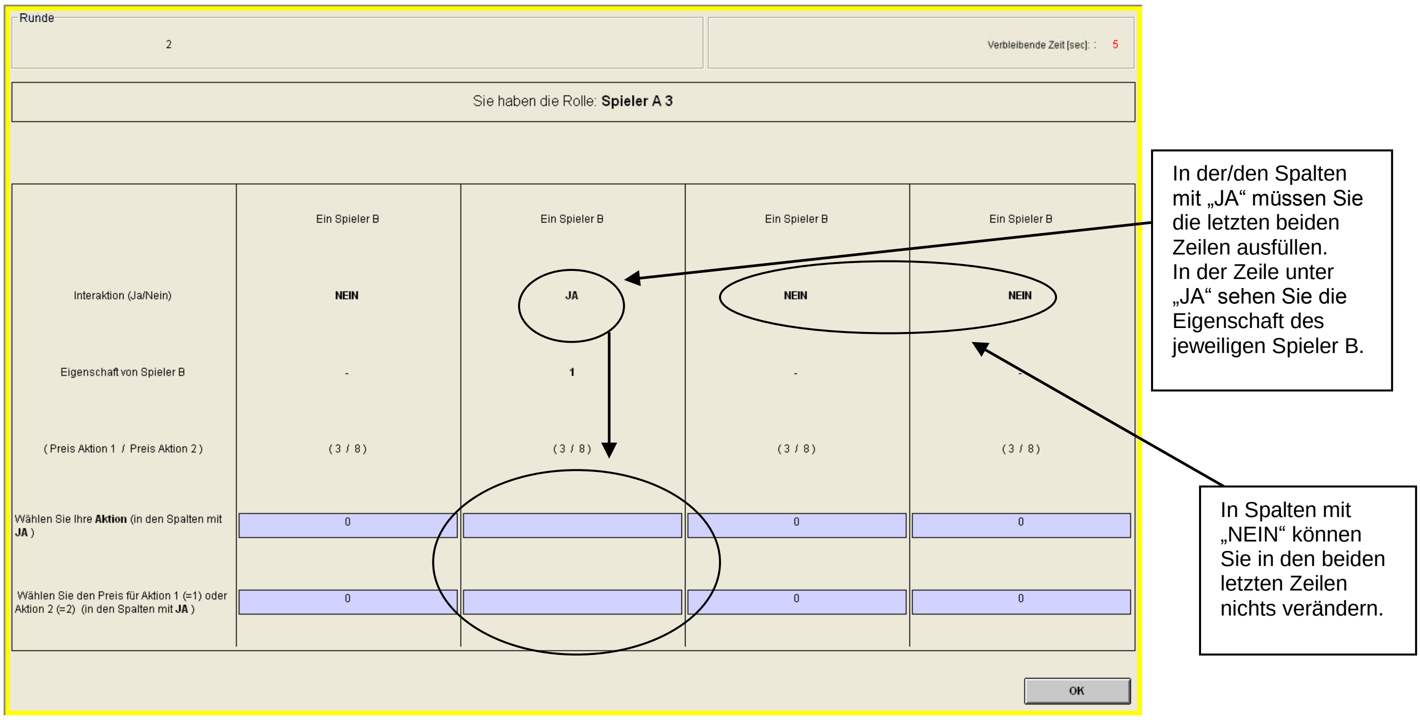
Spieler B erhält **10 Punkte**, wenn die von Spieler A gewählte **Aktion ausreichend** ist. **Spieler B** erhält **0 Punkte**, wenn die von Spieler A gewählte **Aktion nicht ausreichend** ist. In beiden Fällen ist noch der entsprechende Preis zu bezahlen (siehe unten bei „Auszahlungen“).

Spieler B wird zu **keiner** Zeit auf dem Computerbildschirm darüber informiert, ob er/sie in einer Runde die Eigenschaft 1 oder die Eigenschaft 2 hatte bzw. welche Aktion Spieler A gewählt hat.

Entscheidung 4

Spieler A **verlangt** von jedem Spieler B, der mit ihm interagiert, den in Entscheidung 1 festgelegten **Preis** für eine der beiden Aktionen. Dabei **muss** der verlangte Preis **nicht** gleich dem Preis der in Entscheidung 3 gewählten Aktion sein, sondern es kann auch der Preis der anderen Aktion sein. Auch kann Spieler A von unterschiedlichen Spielern B (wenn mehrere Spieler B mit ihm interagieren) unterschiedliche Preise verlangen.

Im Folgenden sehen Sie einen exemplarischen Bildschirm für die Entscheidungen 3 und 4. Jeder Spieler A erfährt für jeden der 4 zufällig gereihten Spieler B, ob der betreffende Spieler B mit ihm interagieren möchte oder nicht (erste Zeile). Falls „JA", dann steht in der entsprechenden Spalte die Eigenschaft von Spieler B. Darunter sind zur Wiederholung die Preise angegeben, die Spieler A in Entscheidung 1 festgesetzt hat.
Die beiden letzten Zeilen sind dann für jene Spalten auszufüllen, in denen bei Interaktion „JA" steht. In der vorletzten Zeile muss für jeden Spieler B eine Aktion gewählt werden (1 oder 2) und in der letzten Zeile muss angegeben werden, welchen Preis Spieler A verlangen möchte (1 steht für den Preis für die Aktion 1; 2 steht für den Preis für die Aktion 2). Auf dem Beispielsbildschirm wollte ein Spieler B mit dem betrachteten Spieler A interagieren und für diese Spalten muss Spieler A seine Entscheidungen eingeben (d.h. die „0"-en ersetzen).



<u>Auszahlungen</u>
Keine Interaktion
Wenn **Spieler B** in Entscheidung 2 mit keinem Spieler A interagiert (*Entscheidung „**Nein**“ für alle 4 Spieler A*), dann erhält er in dieser Runde **1,6 Punkte**.
Wenn kein Spieler B mit einem bestimmten Spieler A interagiert, dann erhält dieser **Spieler A** in dieser Runde auch **1,6 Punkte**.

Ansonsten (*Entscheidung „Ja“ von Spieler B*) sind die Auszahlungen wie folgt:

Interaktion
Spieler A erhält für jeden Spieler B, der mit ihm interagiert, seinen in Entscheidung 4 gewählten **Preis** (in Punkten) **abzüglich** der **Kosten** (siehe Seite 1 unten) für die in Entscheidung 3 gewählte Aktion. D.h. die Auszahlung eines Spielers A setzt sich aus allen Interaktionen zusammen, die ein Spieler A in einer bestimmten Runde hat.

Für **Spieler B** hängt die Auszahlung davon ab, ob die vom betreffenden Spieler A in Entscheidung 3 gewählte Aktion ausreichend war.

a) Die Aktion von Spieler A war ausreichend. **Spieler B** erhält **10 Punkte abzüglich** des in Entscheidung 4 verlangten **Preises**.
b) Die Aktion von Spieler A war nicht ausreichend. **Spieler B** muss den in Entscheidung 4 verlangten Preis bezahlen.

Zu Beginn des Experiments erhalten Sie eine **Anfangsausstattung von 6 Punkten**. Außerdem erhalten Sie durch das Beantworten der Kontrollfragen 2 Euro (entspricht **8 Punkten**). Aus diesen Anfangsausstattungen können Sie auch mögliche Verluste in einzelnen Runden bezahlen. Verluste sind aber auch durch Gewinne aus anderen Runden ausgleichbar. Sollten Sie am Ende des Experiments in Summe einen Verlust gemacht haben, müssen Sie diesen Verlust an den Experimentleiter bezahlen. Mit Ihrer Teilnahme am Experiment erklären Sie sich mit dieser Bedingung einverstanden. Beachten Sie aber bitte, dass es in diesem Experiment **immer** eine Möglichkeit gibt, Verluste mit Sicherheit zu vermeiden.

Für die Auszahlung werden die Anfangsausstattungen und die Gewinne aller Runden zusammengezählt und mit folgendem Umrechnungskurs am Ende des Experiments in bares Geld umgetauscht:

1 Punkt = 25 Euro-Cent
(d.h. 4 Punkte = 1 Euro).

5 Concluding Remarks

This thesis analyzes the impact of expert and market characteristics on the expert's incentive to defraud his customer. In chapter 2, we make use of a field study in order to investigate the drivers of the expert's incentive to overcharge. Focusing on corporate garages, we find that a high level of competition reduces the incentive to overcharge. This might be due to fact that a high level of competition implies low customer search costs for a second opinion. Thus, the probability of a customer to search for a different expert when receiving the diagnosis of a major problem is high. Experts in a critical financial situation overcharge more often than experts in a solid financial situation and thereby gamble for resurrection. Also, experts facing low reputational concerns overcharge more often than those with high reputational concerns. The reason is that experts with low reputational concerns only lose current but no future profit in case the customer visits a different expert after receiving a major diagnosis. Furthermore, the study shows that a high expert competence comes along with a low incentive to overcharge. The study presented in chapter two thereby delivers practical customer implications. Customers should avoid garages that have a low reputational concern and those in a critical financial situation. Instead, customers should search for garages in competitive environments. If customers are able to get a hint of the car repair shop's competence, they should not only choose the better competence because they want their problems to be fixed but also because those garages face a lower incentive to overcharge.

While the car inspection market is characterized by certain price ranges that the garages have to meet, another promising avenue of field research is the analysis of price changes on the experts' incentive to defraud. In the German health care market, there have been two exogenous price shocks within the last fifteen years. The first was a considerable decline in prices when introducing the budget for outpatient treatments while prices remained constant for inpatient treatments. The second change in prices was when Diagnosis Related Groups (DRGs) were introduced. Inpatient treatments are remunerated by a lump sum instead by a fee for service ever since while outpatient remuneration did not change. It would be interesting to exploit these natural experiments in order to analyze how the price changes impacted the physicians' treatment, charging, and defrauding behavior.

Chapter 3 introduces a model where experts (physicians) and customers (patients) both have other-regarding preferences. It is shown that the level of overcharging in the market does not necessarily decrease but may even increase when patients' trust costs increase—depending on the market equilibrium. Thus, the implications of the model are twofold: sensitizing patients for overcharging lowers the level of fraud if the market currently experiences a low level of fraud but it increases fraud if the level of fraud is already high. A more morally focused education of physicians—leading to an increase in the physicians' trust costs—will always decrease the number of searches for a second opinion and therefore increase welfare. Whereas in the model it is assumed that each physician is working self-employed and independently, there has been an increasing trend towards physician networks. Previous literature has mainly focused on competitive effects of such physician networks (see, e. g., Haas-Wilson and Gaynor, 1998; Sacher and Silvia, 1998). It might be interesting to study—theoretically and empirically—how physician networks impact physicians' defrauding behavior. One could expect that overtreatment is much more present within than outside a physician network as referrals within a network can easily be rewarded by referring patients back.

Chapter 4 provides theoretical predictions and experimental evidence on the question of how price competition and different reputation mechanisms impact the expert's incentive to defraud. We show that the level of undertreatment and the level of overcharging are significantly lower under fixed than under flexible prices. This is due to the fact that the price competition undermines the expert's incentive to build up a good reputation by not undertreating if prices are flexible. Experts who undertreated in previous periods try to balance their bad reputation by offering low prices. The lower prices in turn make it less attractive for the expert to treat sufficiently and to charge honestly. Interestingly, we find indicative evidence that under flexible prices, there is more expert fraud if customers' treatment histories are public compared to when customers' histories are private. The reason is that price competition is even more intense under public than under private histories. This again reduces the expert's incentive to treat his customers sufficiently and to charge them honestly. Our results are particularly interesting as Dulleck et al. (2012) find opposite effects when varying the price system in markets where reputation building is not possible: in such a market, price setting is the only way for experts to signal honest treatment. Dulleck et al. (2012) show that the level of undertreatment is lower and efficiency is higher under flexible than under fixed prices. Combining the results of the two studies delivers two important implications for regulators: Firstly, more customer information does not necessarily lead to a lower level of fraud nor does it necessarily lead to a higher efficiency. Secondly, whether flexible or fixed prices induce a lower level of fraud crucially depends upon which reputation mechanism is in place.

This thesis provides important insights into the impact of market and expert characteristics on the expert's incentive to defraud. A major assumption for the models here and in all other contributions on credence goods is, however, that customers know the costs of the treatments.[56] Considering the fact that customers do not even know which treatment they need, this assumption seems to be demanding. Yet, relaxing the cost assumption makes it impossible for customers to infer from experts' mark-ups to experts' incentive to overcharge and to overtreat. It remains to be

[56]To the best of my knowledge, the only exception is Hilger (2011).

investigated how the characteristics of a credence goods market shape the expert's fraudulent behavior under such conditions.

Bibliography

Ahlert, M., Felderer, S., Vogt, B., 2008. How Economists and Physicians Trade Off Efficiency and Equity in Medically and Neutrally Framed Allocation Problems? Working Paper.

Ai, C., Norton, E. C., 2003. Interaction Terms in Logit and Probit Models. Economics Letters 80 (1), 123–129.

Alger, I., Salanié, F., 2006. A Theory of Fraud and Overtreatment in Experts Markets. Journal of Economics & Management Strategy 15 (4), 853–881.

Ariely, D., Mazar, N., 2006. Dishonesty in Everyday Life and its Policy Implications. Journal of Public Policy and Marketing 25 (1), 117–126.

Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2013. What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods. Forthcoming in: Review of Economic Studies.

Beck, A., Kerschbamer, R., Qiu, J., Sutter, M., 2011. Guilt from Promise-Breaking and Trust in Markets for Expert Services: Theory and Experiment. Working Paper.

Bennett, V., Pierce, L., Snyder, J., Toffel, M., 2013. Customer-Driven Misconduct: How Competition Corrupts Business Practices. Forthcoming: Management Science.

Boone, J., 2008. A New Way to Measure Competition. The Economic Journal 118 (531), 1245–1261.

Consumer Federation of America, National Association of Consumer Agency Administrators, North American Consumer Protection Investigators, 2011. 2010 Consumer Complaint Survey Report.

Darby, M. R., Karni, E., 1973. Free Competition and the Optimal Amount of Fraud. Journal of Law and Economics 16 (1), 67–88.

Dulleck, U., Johnston, D., Kerschbamer, R., Sutter, M., 2012. The Good, the Bad and the Naive: Do Fair Prices Signal Good Types or Do They Induce Good Behaviour? Working Paper.

Dulleck, U., Kerschbamer, R., 2006. On Doctor, Mechanics, and Computer Specialists: The Economics of Credence Goods. Journal of Economic Literature 44 (1), 5–42.

Dulleck, U., Kerschbamer, R., 2007. Price Discrimination via the Choice of Distribution Channels. Working Paper.

Dulleck, U., Kerschbamer, R., 2009. Experts vs. Discounters: Consumer Free-Riding and Experts Withholding Advice in Markets for Credence Goods. International Journal of Industrial Organization 27 (1), 15–23.

Dulleck, U., Kerschbamer, R., Sutter, M., 2009. The Impact of Distributional Preferences on (Experimental) Markets for Expert Services. Working Paper.

Dulleck, U., Kerschbamer, R., Sutter, M., 2011. The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. American Economic Review 101 (2), 526–555.

Ely, J., Fudenberg, D., Levine, D. K., 2008. When is Reputation Bad? Games and Economic Behavior 63 (2), 498–526.

Ely, J. C., Välimäki, J., 2003. Bad Reputation. Quarterly Journal of Economics 118 (3), 785–814.

Emons, W., 1997. Credence Goods and Fraudulent Experts. RAND Journal of Economics 28 (1), 107–119.

Emons, W., 2001. Credence Goods Monopolists. International Journal of Industrial Organization 19 (3-4), 375–389.

Federal Bureau of Investigation, 2007. Financial Crimes Report to the Public: Fiscal Year 2007.

Fehr, E., Gächter, S., 2002. Altruistic Punishment in Humans. Nature 415 (6868), 137–140.

Firth, D., 1993. Bias Reduction of Maximum Likelihood Estimates. Biometrika 80 (1), 27–38.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. Experimental Economics 10 (2), 171–178.

Fong, Y.-f., 2005. When Do Experts Cheat and Whom Do They Target? RAND Journal of Economics 36 (1), 113–130.

George, R. B., McKeen, D., Columb, M. O., Habib, A. S., 2010. Up-Down Determination of the 90% Effective Dose of Phenylephrine for the Treatment of Spinal Anesthesia-Induced Hypotension in Parturients Undergoing Cesarean Delivery. Anesthesia & Analgesia 110 (1), 154–158.

Gneezy, U., 2005. Deception: The Role of Consequences. American Economic Review 95 (1), 384–394.

Greiner, B., 2004. The Online Recruitment System ORSEE 2.0—A Guide for the Organization of Experiments in Economics. Working Paper.

Gruber, J., Owings, M., 1996. Physician Financial Incentives and Cesarean Section Delivery. RAND Journal of Economics 27 (1), 99–123.

Haas-Wilson, D., Gaynor, M., 1998. Physician Networks and Their Implications for Competition in Health Care Markets. Health Economics 7 (2), 179–182.

Heinze, G., 2006. A Comparative Investigation of Methods for Logistic Regression with Separated or Nearly Separated Data. Statistics in Medicine 25 (24), 4216–4226.

Hennig-Schmidt, H., Wiesen, D., 2010. Are Prospective Physicians Different? Evidence from an Experiment on a Payment System Variation. Working Paper.

Hilger, N. G., 2011. Why Don't People Trust Experts? Working Paper.

Hirschman, A. O., 1964. The Paternity of an Index. American Economic Review 54 (5), 761–762.

Hubbard, T. N., 1998. An Empirical Examination of Moral Hazard in the Vehicle Inspection Market. RAND Journal of Economics 29 (2), 406–426.

Huber, P. J., 1967. The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions. In: Le Cam, L. M., Neyman, J. (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1. University of California Press, pp. 221–233.

Iizuka, T., 2007. Experts' Agency Problems: Evidence From the Prescription Drug Market in Japan. RAND Journal of Economics 38 (3), 844–862.

Jin, G. Z., Leslie, P., 2009. Reputational Incentives for Restaurant Hygiene. American Economic Journal: Microeconomics 1 (1), 237–267.

Klein, B., Leffler, K. B., 1981. The Role of Market Forces in Assuring Contractual Performance. Journal of Political Economy 89 (4), 615–641.

Kreps, D. M., Wilson, R., 1982. Reputation and Imperfect Information. Journal of Economic Theory 27 (2), 253–279.

Liu, T., 2011. Credence Goods Markets With Conscientious and Selfish Experts. International Economic Review 52 (1), 227–244.

McFadden, D., 1979. Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments. In: Hensher, D. A., Stopher, P. R. (Eds.), Behavioural Travel Modeling. Croom Helm, London, pp. 279–318.

McKelvey, R. D., Zavoina, W., 1975. A Statistical Model for the Analysis of Ordinal Level Dependent Variables. Journal of Mathematical Sociology 4 (1), 103–120.

Milgrom, P., Roberts, J., 1982. Predation, Reputation, and Entry Deterrence. Journal of Economic Theory 27 (2), 280–312.

OECD, 2011. Health at a Glance 2011: OECD Indicators. OECD Publishing, Danvers.

Park, I.-U., 2005. Cheap-Talk Referrals of Differentiated Experts in Repeated Relationships. RAND Journal of Economics 36 (2), 391–411.

Pesendorfer, W., Wolinsky, A., 2003. Second Opinions and Price Competition: Inefficiency in the Market for Expert Advice. Review of Economic Studies 70 (2), 417–437.

Pieper, A., Thurnherr, U., 1998. Angewandte Ethik: Eine Einführung. Beck, München.

Pike, C., 2010. An Empirical Analysis of the Effects of GP Competition. Working Paper.

Richardson, H., 1999. The Credence Good Problem and the Organization of Health Care Markets. Working Paper.

Rogers, W. H., 1993. Regression of Standard Errors in Clustered Samples. Stata Technical Bulletin 3 (13), 19–23.

Sacher, S., Silvia, L., 1998. Physician Networks, Integration and Efficiency. Working Paper.

Schneider, H. S., 2012. Agency Problems and Reputation in Expert Services: Evidence From Auto Repair. Forthcoming: Journal of Industrial Economics.

Shapiro, C., 1983. Premiums for High Quality Products as Returns to Reputations. Quarterly Journal of Economics 98 (4), 659–679.

Statistisches Bundesamt (Ed.), 2012. Gesundheit: Ausgaben: 2010. Wiesbaden.

Sülzle, K., Wambach, A., 2005. Insurance in a Market for Credence Goods. Journal of Risk and Insurance 72 (1), 159–176.

Taylor, C. R., 1995. The Economics of Breakdowns, Checkups, and Cures. Journal of Political Economy 103 (1), 53–74.

Tibandebage, P., Mackintosh, M., 2005. The Market Shaping of Charges, Trust and Abuse: Health Care Transactions in Tanzania. Social Science & Medicine 61 (7), 1385–1395.

Titus, R. M., Heinzelmann, F., Boyle, J. M., 1995. Victimization of Persons by Fraud. Crime & Delinquency 41 (1), 54–72.

Transparency International - Deutschland e.V. (Ed.), 2008. Transparenzmängel, Korruption und Betrug im deutschen Gesundheitswesen: Kontrolle und Prävention als gesellschaftliche Aufgabe. Berlin.

White, H., 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica, 817–838.

Wolinsky, A., 1993. Competition in a Market for Informed Experts' Services. RAND Journal of Economics 24 (3), 380–398.

Wolinsky, A., 1995. Competition in Markets for Credence Goods. Journal of Institutional and Theoretical Economics 151 (1), 117–131.

Wooldridge, J. M., 2009. Introductory Econometrics: A Modern Approach. Cengage Learning Services, Mason.

Zentralverband deutsches Kraftfahrtzeuggewerbe (Ed.), 2012. Zahlen & Fakten 2011: Ausgabe 2012. Self-published, Bonn.

Zorn, C., 2005. A Solution to Separation in Binary Response Models. Political Analysis 13 (2), 157–170.

Curriculum Vitae

Personal Details

Place of birth	Muenster, Germany
Date of birth	May 20, 1984
Nationality	German

Education

10/2009–01/2013	Ph.D. student at the Cologne Graduate School in Management, Economics and Social Sciences University of Cologne, Germany Supervisor: Prof. A. Wambach, Ph.D.
10/2011–01/2012	Visiting Ph.D. student Department of Public Finance (Prof. Dr. M. Sutter) University of Innsbruck, Austria
07/2007–12/2007	Visiting student University of Queensland, Brisbane, Australia
10/2004–12/2009	Studies in Mathematics & Social Sciences Universities of Muenster and Cologne, Germany *Erstes Staatsexamen*
10/2003–03/2009	Studies in Economics Universities of Muenster and Cologne, Germany *Diplom*
08/1994–06/2003	Ratsgymnasium, Muenster, Germany *Abitur*

Cologne, January 2013