RUHR
ECONOMIC PAPERS

Jeannette Brosig-Koch
Burkhard Hehenkamp
Johanna Kokot

The Effects of Competition on Medical Service Provision

UNIVERSITÄT DUISBURG ESSEN

#647

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

RWI Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #647

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-753-3

Ruhr Economic Papers #647

Jeannette Brosig-Koch, Burkhard Hehenkamp, and Johanna Kokot

The Effects of Competition on Medical Service Provision

UNIVERSITÄT DUISBURG ESSEN

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

Das RWI wird vom Bund und vom Land Nordrhein-Westfalen gefördert.

http://dx.doi.org/10.4419/86788753
ISSN 1864-4872 (online)
ISBN 978-3-86788-753-3

Jeannette Brosig-Koch, Burkhard Hehenkamp, and Johanna Kokot[1]

The Effects of Competition on Medical Service Provision

Abstract

We explore how competition between physicians affects medical service provision. Previous research has shown that, without competition, physicians deviate from patient-optimal treatment under payment systems like capitation and fee-for-service. While competition might reduce these distortions, physicians usually interact with each other repeatedly over time and only a fraction of patients switches providers at all. Both patterns might prevent competition to work in the desired direction. To analyze the behavioral effects of competition, we develop a theoretical benchmark which is then tested in a controlled laboratory experiment. Experimental conditions vary physician payment and patient characteristics. Real patients benefit from treatment decisions made in the experiment. Our results reveal that, in line with the theoretical prediction, introducing competition can reduce overprovision and underprovision, respectively. The observed effects depend on patient characteristics and the payment system, though. Tacit collusion is observed and particularly pronounced with fee-for-service payment, but it appears to be less frequent than in related experimental research on price competition.

JEL Classification: I11, D43, C91, C72

Keywords: Physician competition; fee-for-service; capitation; laboratory experiment

October 2016

---

1 Jeannette Brosig-Koch, UDE and CINCH; Burkhard Hehenkamp, University of Paderborn; Johanna Kokot, UDE. *- Financial support provided by the German Research Foundation (DFG; BR 2346/2-1) is gratefully acknowledged.* *– All correspondence to: Jeannette Brosig-Koch, Faculty of Economics and Business Administration & Health Economics Research Center CINCH, Berliner Platz 6-8, 45127 Essen, Germany, e-mail: jeannette.brosig@uni-due.de*

1 Introduction

Starting with the seminal papers by Ellis and McGuire (1986, 1990) and Ellis (1998), an extensive literature has investigated the extent to which physician payment systems like capitation and fee-for-service lead to deviations from patient-optimal medical treatment (for an overview see, e.g., Iversen and Lurås, 2006). This literature includes both theoretical and empirical contributions, the latter of which mostly contain field evidence. While capitation payment embeds an incentive to provide fewer medical services than would be optimal for the patient, fee-for-service payment induces physicians to supply more than the patient's optimal level of medical service (see, e.g., Ellis and McGuire, 1986).

Field evidence on the relationship between physician payment and medical treatment decisions is rather mixed. Some studies observe that physicians respond to payment incentives (e.g., Clemens and Gottlieb, 2014, Davidson et al., 1992, Devlin and Sarma, 2008, Gaynor and Gertler, 1995). Others do not find a strong link, however (e.g., Hutchinson et al., 1996, Hurley and Labelle, 1995, Grytten and Sørensen, 2001). Field research often struggles with simultaneous variations of more than one component of the payment system or potential selection biases regarding patient characteristics, which make causal inferences on the direction and strength of an effect rather difficult (e.g., Gosden et al., 2001, Falk and Heckman, 2009).

In recent years, research in health economics has started to use economic experiments to test the behavioral effects of physician payment under controlled laboratory conditions (e.g., Hennig-Schmidt et al., 2011, Green, 2014, Brosig-Koch et al., 2015a, b). While there is an active debate on the extent to which insights from the laboratory can be generalized to the field (see, e.g., the discussion in Levitt and List, 2007, or the recent findings by Herbst and Maas, 2015), laboratory experiments typically serve to complement field research as they allow for a higher internal validity. In the laboratory *ceteris paribus* changes of parameters can be implemented and their effects on individual behavior can be directly observed. External aspects like patients' health status can be isolated and, if behavior changes, this variation can be attributed to the modified parameter (e.g., the payment system). Accordingly, laboratory experiments provide a suitable tool to test health economic models.[1]

Experimental research on physician payment has revealed that monetary incentives affect medical treatment. In line with the theoretical prediction, patients receive significantly more medical services under fee-for-service payment than under capitation payment. This holds

[1] An elaborate discussion of the relationship between economic theory and experimental economics is included in Part II of the Handbook of Experimental Economic Methodology (Fréchette and Schotter, 2015).

true independently of the subject pool, i.e. physicians, medical students, and non-medical students (Brosig-Koch et al., 2015a). Laboratory studies further suggest that medical service provision is not only guided by individual profit, but also by patient health benefit. Based on experimental data provided by Hennig-Schmidt et al. (2011), Godager and Wiesen (2013) explicitly measure the weight subjects attach to patient health benefit. They find that the majority of subjects put a higher weight on patient health benefit than on own profit. This supports the assumption of physician altruism commonly made in the theoretical health economics literature (see, e.g., Ellis and McGuire, 1986, 1990, McGuire, 2000, Chalkley and Malcomson, 1998, Allard et al., 2011, Choné and Ma, 2011). Godager and Wiesen (2013) further observe a substantial heterogeneity in the degree of individual altruism. Similarly, focusing on mixed payment systems, Brosig-Koch et al. (2015b) demonstrate that individual responses can be accounted for by a behavioral model capturing physician altruism. In line with Godager and Wiesen (2013), they find that the weight subjects attach to patients' health benefit differs substantially among subjects. Again, these results do not depend on subjects' medical background.

So far the experimental literature restricts attention to medical service provision that is made in the absence of competition. We contribute to this literature by exploring how competition affects the distortion of medical treatment caused by payment incentives. Based on a model of physician competition, we derive behavioral predictions which are then tested in a controlled laboratory experiment. Motivated by the experimental findings on physician altruism, our model allows for heterogeneity in the weight individuals attach to patients' health benefit.

Theoretical research mostly admits that competition between physicians can reduce the distortion of behavior under certain conditions (see e.g. Allard et al., 2009, or the overview provided by Gaynor and Town in the Handbook of Health Economics, 2012). Merely, Ellis and McGuire (1986) argue that hospital competition for physicians will strengthen the distortionary effect of payment systems. Notice, however, that this issue does not arise when independent physicians compete for patients.[2] Gravelle and Masiero (2000) and Karlsson (2007) study factors that presumably alleviate the effect of introducing competition between general practitioners (GPs) under a capitation system. Both papers consider a model of horizontally and vertically differentiated GPs competing for patients who are imperfectly informed about quality and face costs of switching providers. To the extent that quality remains uncertain

[2] According to Ellis and McGuire, hospitals primarily compete for physicians rather than for patients directly and they compete for physicians with e.g. a lower (higher) weight on patient benefit under a prospective (cost-based) payment system. As a consequence, intensified *competition for physicians* will tend to strengthen the distortionary impact originating from the payment systems.

even after medical treatment, service provision exhibits credence good characteristics in these studies.[3] Gravelle and Masiero (2000) devote special attention to the problem of supplier induced demand in a model with boundedly rational patients. They show that uncertainty in the assessment of quality lowers quality levels and weakens the incentive effects that result from an increase in the capitation fee. Welfare increases with reduced uncertainty at a diminishing rate and increases with reduced switching costs at an increasing rate. Karlsson (2007) investigates a model where rational patients remain uncertain of the quality even after having consumed it. Quality varies due to GP heterogeneity. Karlsson finds that under such circumstances an increase in the capitation fee has direct effects on quality, but these will be (partially) offset by changes in the search activity of patients. In a recent experimental study, Huck et al. (2014) examine the provision of a credence good, employing a 2x2-design with insurance and competition. Compared to their baseline treatment, insurance leads to higher consultation rates and higher overtreatment (two-sided moral hazard). The adverse effects of insurance are mitigated when competition is combined with insurance.

Field evidence on the effects of physician competition is scarce and results are rather mixed. Pike (2010) investigates the relationship between GP's quality of medical care (number of referrals to specialists, patient satisfaction) and the degree of competition (number of nearby rivals) in England. He reports that more competition is correlated with a higher level of quality. Dunn and Shapiro (2015) focus on the impact of competition on the quantity and type of health services provided by cardiologists. They observe that, with fee-for-service payment, a higher market concentration increases the use of cardiac catheterization, but decreases the probability of a less invasive diagnostic test being performed. Moreover, a higher concentration leads to fewer readmissions, but does not affect mortality. Iversen and Ma (2011) use Norwegian data of GP radiology referrals to study the relationship between competition and the number of referrals. In line with their model, they find that competition leads to a higher number of referrals. Godager et al. (2015) re-examine the effect of competition on GP referrals in Norway and include some additional controls. According to their results, competition has no or only a small positive effect on the number of referrals.

In this study, we use a controlled laboratory experiment to test the effects of competition on medical treatment decisions. Theoretical predictions are based on a model which allows for

[3] There exists a related literature on markets for credence goods, which is nicely summarized and extended in Dulleck and Kerschbamer (2006). However, this literature differs from our approach in that it presumes that prices are set by providers and that providers maximize profit. In contrast, the literature on competition in health care typically regards prices as regulatory instruments taken as given by providers. Moreover, in this literature providers are usually assumed to be altruistic to some degree, maximizing a weighted average of own profit and patient's utility.

different degrees of physician altruism (modeled as a weighted average of physician profit and patient benefit). In our set-up, two subjects take the role of physicians and repeatedly interact with each other over 20 rounds. In every round there are four patients to be treated by the two physicians. Two patients are permanently allocated to the same physician with each physician treating one of them. The other two patients are recurrently reassigned to the physician providing the highest health benefit. If both physicians provide identical health benefit, patients split equally between the two physicians. In each round, the two physicians simultaneously decide on their level of medical treatment. Each pair of decisions jointly determines the physicians' profit, which depends on the payment system and the number of patients treated. Experimental conditions systematically vary patient health outcomes and physician incentives. This allows to isolate the effects of competition and, in particular, to control for potentially important factors like patient characteristics and payment systems.

Our assumptions regarding patient behavior serve to reflect the observation that people choose their doctor largely on the basis of convenience and/or some form of quality (e.g., Salisbury, 1989, Dixon et al., 1997, Biørn and Godager, 2010). Given that quality has several dimensions, which vary in their degree of observability, our set-up captures situations where a patient chooses a doctor largely on the basis of observable quality dimensions (see also the research on public quality reports, e.g. Dranove and Jin, 2010), but also situations where the patient's decision is informed by an expert such as a GP who acts as a gate-keeper to specialist services (see e.g. Brekke et al, 2007, Beukers et al, 2014).

The experiment is designed such that a trade-off arises between patient-optimal and profit-maximal treatment decisions in the absence of competition and that introducing competition yields a unique theoretical prediction, where both physicians choose the patient-optimal treatment level. This holds for the stage game and for the finitely repeated game. The theoretical prediction of patient-optimal treatment is robust against introducing altruistic preferences on the part of the physicians, both for the case of commonly known degrees of altruism and when a physician's degree of altruism represents her private information. Theoretically, our framework leaves no scope for collusive behavior (i.e., deviations from the patient optimum in the direction of the joint profit maximum). Previous experimental research on finitely repeated games reveals that collusion can still be observed (at least for earlier rounds of the interaction; see, e.g., Potters and Suetens, 2013). Our experiment allows to test whether this result holds also for our medical decision setting in which real patients benefit from treatment decisions and, if so, whether it depends on patient characteristics and physician payment.

The paper is organized as follows. In section 2, we describe the experimental design and procedure. Section 3 presents the theoretical predictions and results and section 4 our experimental findings. In section 5, we sum up and conclude.

2 Experimental Set-up

2.1 Design

In all experimental ***conditions with competition***, subjects face the following decision situation: Taking the role of physicians, subjects are randomly and anonymously matched in pairs, which remain fixed over the 20 rounds of the experiment. In each round, each of the two subjects i in each pair simultaneously and independently decide on the level of medical treatment $q_i \in Q \coloneqq \{0, 1, \dots, 10\}$, which is then applied to all of her patients in that round.[4] Any decision on q_i has three effects: It determines the health benefit of patients treated by this physician, $B(q_i)$, it determines the physician's profit per patient treated, $\pi(q_i)$, and it affects the number of patients treated, $n_i(q_1, q_2)$. Each pair of subjects is faced with four patients who exhibit identical health characteristics. Patients only differ regarding their mobility: *Regular patients* always visit the same subject while *undecided patients* visit the subject providing the highest patient benefit. In case of a tie, undecided patients split up evenly. In our set-up, there is one regular patient assigned to each subject. Accordingly, depending on the treatment choices made by both subjects, a subject treats at least one and at most three patients. To sum up, we assume

$$n_i(q_1, q_2) = \begin{cases} 1 & \text{if } B(q_i) < B(q_j) \\ 2 & \text{if } B(q_i) = B(q_j) \\ 3 & \text{if } B(q_i) > B(q_j) \end{cases}$$

for $i, j = 1{,}2$ and $j \neq i$.

For each of our competition conditions we employ a baseline ***condition without competition*** where we set $n_i(q_1, q_2)$=1 and where profit and benefit functions remain identical to the respective competition condition. Taking the role of a physician, each subject independently decides on the level of medical treatment for one regular patient per round. As in the competition conditions, decisions are repeated over 20 rounds.

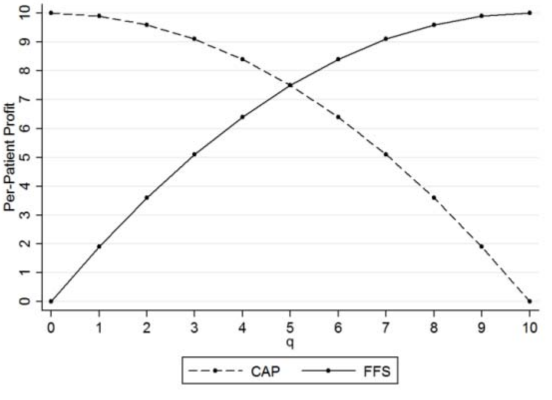
In the following we present details about the design of physicians' profits and patients' benefits. The design is based on Brosig-Koch et al. (2015a, b).

[4] The level of medical treatment q can be interpreted as investment in medical equipment (e.g., in new technologies or in the development of new skills).

Profit

For each patient treated, a subject receives a remuneration $R(q)$ and incurs a cost $c(q)=0.1q^2$.[5] There are two types of remuneration tested in the experiment – fee-for-service (FFS) and capitation (CAP). In CAP, each subject receives a lump-sum payment per patient of 10, i.e. $R^{CAP}(q) = 10$. In FFS, the remuneration increases with the treatment level, i.e. $R^{FSS}(q) = 2q$. Accordingly, a subject's *profit per patient* is $\pi^{CAP}(q) = 10 - 0.1q^2$ in CAP and $\pi^{FSS}(q) = 2q - 0.1q^2$ in FFS. Observe that the profit per patient is symmetric between CAP and FFS in that $\pi^{CAP}(q) = \pi^{FSS}(10 - q)$ for all $q \in Q$. Figure 1 illustrates the profit per patient for both experimental conditions, CAP and FFS.

Figure 1: Per-patient profit in CAP and in FFS



Subject i's *total profit* is given by the number of patients treated times the profit per patient, i.e. $\pi_i(q_1, q_2) = n_i(q_1, q_2)\pi(q_i)$, where $\pi(q) \in \{\pi^{CAP}(q), \pi^{FSS}(q)\}$ depends on the experimental condition.

Patient Benefit

Each level of medical treatment q results in a patient benefit $B(q)$. We distinguish two patient types: For patients with a high severity of illness (H) the patient-optimal level is $q^* = q^H = 7$. For patients with a low severity of illness (L) the patient-optimal level of medical treatment is $q^* = q^L = 3$. The patient-optimal level serves as a benchmark for identifying the extent of overprovision and underprovision, respectively. The maximum patient benefit is 10 for both patient types. Accordingly, we set

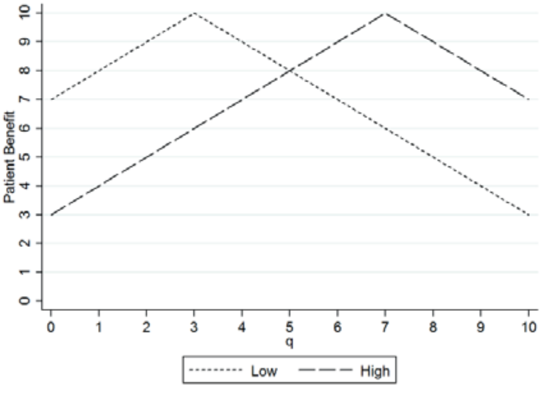
[5] Convex cost functions are used in several theoretical models describing physician behavior (e.g., Ma, 1994, and Choné and Ma, 2011).

$$B^L(q) = 10 - |q^L - q| \qquad \text{where } q^L \coloneqq 3 \text{ in L},$$
$$B^H(q) = 10 - |q^H - q| \qquad \text{where } q^H \coloneqq 7 \text{ in H}.$$

Figure 2 depicts the patient benefit for the two patient types H and L. Observe that the patient benefit $B(\cdot)$ is concave, it is characterized by a unique global maximum, and it is mirror-symmetric at this maximum.[6] The symmetric design of profits and benefits allows to directly compare behavior between the two symmetric payment systems (i.e., it allows to test whether incentives to underprovide in CAP are equally effective as incentives to overprovide in FFS).

The patient benefit $B(q)$ is given in monetary terms and is known to subjects. While no subject takes the role of patients, real patients outside the laboratory benefit from subjects' treatment decision. In particular, subjects are informed that the monetary value of the total patient benefit is transferred to an organization (*Christoffel Blindenmission*) which cares for real patients with eye cataract (see also Section 2.2 below). So subjects' decisions affect real patients' health.

Figure 2: Patient benefit for H and L



2.2 Experimental Protocol

We conducted the computerized experiment at the Essen Laboratory for Experimental Economics at the University of Duisburg-Essen, Germany. The experiment was programmed with *z*-Tree (Fischbacher, 2007). 178 student participants were recruited using the online recruiting system ORSEE (Greiner, 2004).[7] In total, we employed four different competition conditions and four different no-competition conditions varying the payment system (FFS vs. CAP) and

[6] Symmetry at the maximum refers to $B^j(q^j - q) = B^j(q^j + q)$ for $j \in \{L, H\}$.

[7] As Brosig-Koch et al. (2015a, b) do not find qualitatively different responses to CAP and FFS incentives between medical students and students from other fields, our study is based on a conventional student subject pool.

the patient type (H vs. L). Table 1 provides an overview of the experimental conditions examined in this study. Matching subjects in pairs in our competition conditions, we generated 12 (11) independent observations per session in these conditions. Of the 178 participants in the experiment, 96 were male and 82 were female.

Table 1: Participants per experimental conditions

	Condition	# Participants
No-competition	CAP_L(NC)	19
	CAP_H(NC)	20
	FFS_L(NC)	22
	FFS_H(NC)	23
Competition	CAP_L(C)	24
	CAP_H(C)	24
	FFS_L(C)	22
	FFS_H(C)	24

Upon arrival, subjects were randomly assigned to cubicles. Subsequently, the instructions were handed out and subjects were given sufficient time to read them. Clarifying questions were answered in private. To check whether subjects understood the set-up, they were given a set of control questions. The experiment started once all subjects had answered the questions correctly. At the beginning of each competition condition, subjects were randomly matched in pairs, which remained fixed over the 20 rounds. A history table summarized all relevant information on the subjects' current and past rounds, i.e. the chosen quantities of medical treatment, the benefit per patient treated, the total profit per round, and – for the competition conditions – the number of patients treated by each of the two subjects (for the full set of instructions, see Appendix A).

Physician profit and patient benefit were given in *Taler*. As this experiment tests repeated interaction, each decision round was payoff-relevant. At the end of the experiment, physician profits and patient benefits were summed up and the amounts in *Taler* were multiplied with an exchange factor of 0.05 [0.08] Euro/Taler in the competition [no-competition] conditions. A session lasted for about 90 [60] minutes. Subjects earned on average Euro 14.84. In total, Euro 2704 were transferred to the *Christoffel Blindenmission*. Since eye cataract surgery costs approximately Euro 30, about 90 real patients could be treated. To ensure a credible transfer, we randomly selected a subject after the experiment to monitor the transfer procedure. This subject had to verify that a correct transfer order was sent to the university's financial department. The monitor and experimenter then walked together to the nearest mailbox and deposited the order in a sealed envelope. The monitor was paid an additional Euro 5 (see, e.g., Hennig-Schmidt et al., 2011, for a similar procedure).

3 Theoretical Predictions and Results

We first examine the decision problem without physician competition. Subsequently, we address the case of competition. In particular, we show that, under competition, the theoretical predictions for our design are robust against different degrees of a physician's altruism and against introducing incomplete information about the other physician's degree of altruism, respectively.

Absence of competition

To allow for altruism, let

$$U(q,\alpha) = (1-\alpha)\pi(q) + \alpha B(q)$$

denote the utility of a physician, who exhibits a degree of altruism $\alpha \in [0,1]$ and chooses a treatment level $q \in Q$. Accordingly, a physician maximizes a weighted average of profit $\pi(q)$ and patient benefit $B(q)$. By symmetry of the design, CAP_L and FFS_L represent mirror images of FFS_H and CAP_H, respectively.

In the absence of competition, a profit maximizing physician ($\alpha = 0$) chooses $q^{CAP} := 0$ in CAP and $q^{FFS} := 10$ in FFS, independently of the patients' type (L vs. H). A purely altruistic physician ($\alpha = 1$) maximizes patient benefit and hence chooses $q^* = q^L$ in L and by $q^* = q^H$ in H, independently of the payment system (CAP vs. FFS). In the intermediate case, $\alpha \in (0,1)$, a trade-off arises between maximizing physician profit and maximizing patient benefit. Correspondingly, a physician chooses $q^{CAP}(\alpha) \in \{0,1,\dots,q^*\}$ and $q^{FSS}(\alpha) \in \{q^*,\dots,9,10\}$. As we have $q^* = q^L$ in L and $q^* = q^H$ in H, the trade-off is more pronounced in conditions FFS_L and CAP_H (and less so in FFS_H and CAP_L).

Competition

Under competition, a physician's *total* profit additionally depends on the number of patients treated, $n_i(q_1, q_2)$, which is in turn *jointly* determined by the choice (q_1, q_2) of the two physicians, i.e. $\pi_i(q_1, q_2) = n_i(q_1, q_2)\,\pi(q_i)$. *Total* patient benefit depends on the number of patients treated as well. We set $B_i(q_1, q_2) = n_i(q_1, q_2)\,B(q_i)$, for $i = 1,2$. As before, profit per patient and hence total profit depend on the payment system (CAP vs. FFS) whereas patient benefit depends on the patient type (L vs. H).

To allow for altruistic preferences, let $\alpha = (\alpha_1, \alpha_2) \in [0,1]^2$ denote the degree of altruism of physicians 1 and 2. Each physician is assumed to maximize utility, which represents a weighted average of total profit and total patient benefit, i.e.

$$U_i(q_1, q_2; \alpha_i) = (1 - \alpha_i)\pi_i(q_1, q_2) + \alpha_i B_i(q_1, q_2), \;\; i = 1,2.$$

All this is assumed to be common knowledge among physicians. The proposition below identifies the Nash equilibrium of this stage game for each of the experimental conditions. In particular, it shows that, independently of the payment system, each physician chooses the patient-optimal level of treatment in equilibrium. The proof of the proposition is instructive in that it reveals that our experimental design is robust against introducing incomplete information about each other physicians' degree of altruism. The analysis is restricted to pure strategy Nash equilibria.

Proposition

Let $(\alpha_1, \alpha_2) \in [0,1]^2$. Then the unique Nash equilibrium is given by $(q_1^*, q_2^*) = (q^*, q^*)$, where q^* represents the patients' optimal treatment level, i.e. $q^* = q^L = 3$ in L and $q^* = q^H = 7$ in H. In equilibrium, patient benefit equals 10 per patient in all four experimental conditions, total profit is 10.2 in FFS_L and CAP_H, and it is 18.2 in FFS_H and CAP_L. In any of these cases, the Nash equilibrium is strict.

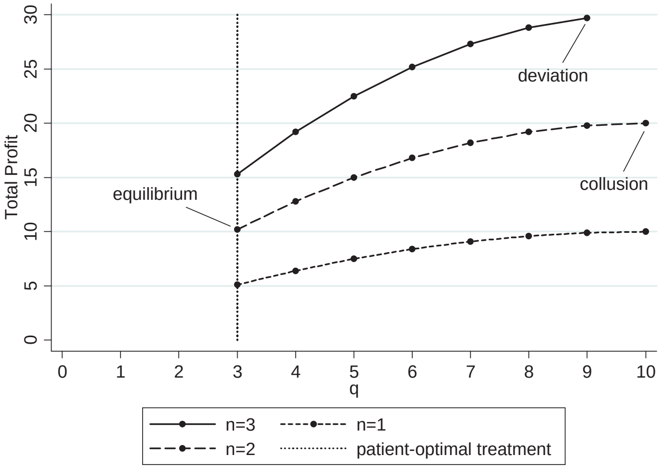
Proof: See Appendix B. □

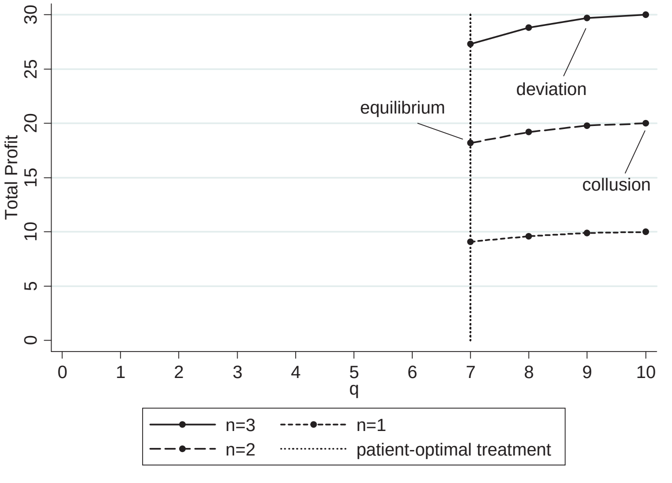
Collusion

Since the stage game is repeated a finite number of times, its unique subgame perfect equilibrium (SPE) involves repeated play of the stage game equilibrium, i.e., on the equilibrium path, the SPE actions coincide with the patient-optimal treatment decisions. Theoretically, our framework leaves no scope for collusion. Based on previous experimental evidence on finitely repeated games, however, collusion is to be expected for the earlier rounds of the interaction (see, e.g., Potters and Suetens, 2013).

Typically, the payoff-related key determinants of cooperative/collusive behavior are considered to be (1) the short-run gain from breaking a cooperative/collusive agreement and (2) the long-run loss from the collapse of future cooperation. Figure 3 illustrates these key determinants for conditions FFS_L and FFS_H, respectively.

Figure 3: Incentives for collusion for patient type L and for patient type H in FFS





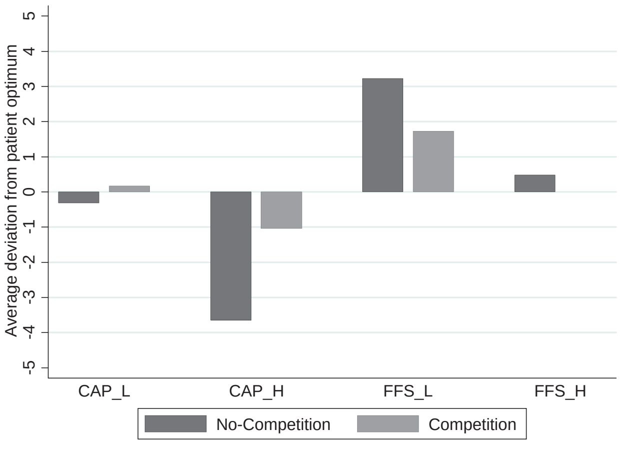
It can be seen that the long-run loss (2) from a collapse of collusion is larger in FFS_L than it is in FFS_H while the short-run gain (1) from breaking the collusive agreement coincides for the two experimental conditions. Therefore, if at all, more collusion should be expected in FFS_L than in FFS_H. Exploiting the symmetry between CAP and FFS, a similar argument demonstrates that more collusion should be expected in CAP_H than in CAP_L.

4 Experimental Results

4.1 First Round Behavior

We start with examining treatment decisions observed in round 1 in the no-competition conditions as, in this round, behavior is not yet affected by learning or experience. We then compare these decisions with those made in round 1 in the competition conditions. Our analysis focuses on subjects' deviations from the patient-optimal treatment level. The lower this deviation is (in absolute terms), the more patients profit from medical treatment. Figure 4 summarizes average deviations from the patient optimum observed in the competition and no-competition conditions.

Figure 4: Average deviation from patient-optimal treatment level



Without competition, we find overprovision in FFS and underprovision in CAP, which is significant for FFS_L(NC) and CAP_H(NC) ($p \leq 0.001$), but not significant for FFS_H(NC) and CAP_L(NC) ($p \geq 0.250$).[8] Observed deviations significantly depend on the patient type ($p \leq 0.001$). The less the patient-optimal treatment level deviates from the profit maximum, the lower is the deviation from this level. As payment systems CAP and FFS as well as patient benefits L and H represent mirror images of each other, we can directly compare behavior between the two payment systems. In line with incentive symmetry, the payment system has no significant effect on treatment decisions (i.e., in absolute terms, average deviations from patient-optimal treatment do not differ significantly from each other; $p \geq 0.467$).

[8] We apply non-parametric tests to the respective averages. For between-subject analyses, we employ exact Mann-Whitney *U* tests. For within-subject analyses we apply exact Wilcoxon matched-pairs signed rank tests. When comparing decisions with predicted treatment levels, we use one-sample Wilcoxon signed rank tests. Throughout the paper, *p*-values are reported from two-sided tests.

Comparing average deviations from patient-optimal treatment between the competition and the no-competition conditions, we observe significantly lower deviations in conditions CAP_L, CAP_H, and FFS_L ($p \leq 0.039$)[9]. Thus, competition reduces the distortions resulting from capitation and fee-for-service incentives. But even with competition, overprovision in FFS_L(C) and underprovision CAP_H(C) are still significant ($p \leq 0.005$).

Also with competition, average deviations from patient-optimal treatment significantly depend on the patient type (CAP_L(C) vs. CAP_H(C), FFS_L(C) vs. FFS_H(C); $p \leq 0.005$). Although incentives are symmetric in CAP_H(C) and FFS_L(C), we observe weakly significantly different treatment decisions in the two conditions ($p=0.058$). There is no significant difference between CAP_L(C) and FFS_H(C) ($p=0.772$). The payment system seems to somewhat affect behavior with competition, but only in conditions in which incentives for collusive behavior are large enough (see section 3).

4.2 Behavior in Later Rounds

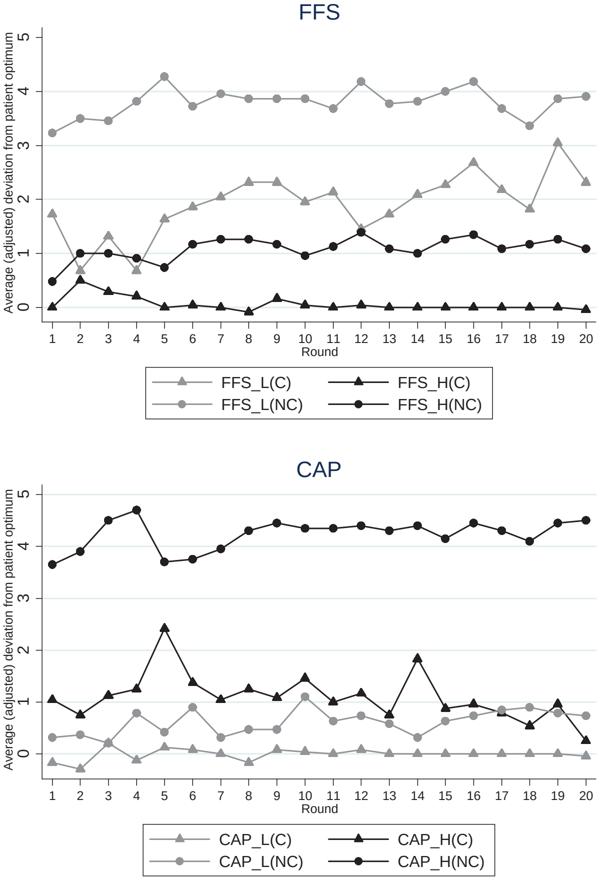
Figure 5 displays the development of average deviations from the patient optimum with competition (lines with triangle) and without competition (lines with circle). To make behavior in the two symmetric payment conditions directly comparable, we use adjusted average deviations in this figure. Adjustments are made such that individual deviations from patient optimum leading to an increase of individual profits are always given a positive sign and individual deviations leading to a decrease of individual profits (i.e., inferior decisions) are always given a negative sign. Due to this adjustment, both underprovision in CAP and overprovision in FFS have a positive sign in Figure 5.

In the absence of competition, we observe quite stable treatment choices in all conditions over the 20 rounds. With competition, average deviations from patient-optimal treatment remain close to zero in FFS_H and CAP_L. In both conditions, 90.1 percent (FFS_H) and 85.8 percent (CAP_L) of the decisions per pair are patient-optimal.[10] Deviations in condition CAP_H(C) seem to converge only slowly to the patient optimum, while there seems to be even an increase of deviations in FFS_L(C). In fact, despite the incentive symmetry of the two payment systems, in the last two rounds we observe a significantly higher adjusted deviation from patient-optimal treatment in FFS_L(C) than in CAP_H(C) ($p \leq 0.050$).

---

[9] Note that the effect of competition in CAP_L is still significant, when excluding the few subjects who choose to *over*provide in the first round of this condition ($p=0.018$).
[10] The difference between the two treatments is not statistically significant ($p=0.360$).

Figure 5: Dynamics of the (adjusted) average deviation from patient optimum



Nevertheless, the frequency of patient-optimal treatment choices per pair is not significantly different between the two conditions (41.7 percent CAP_H, 44.1 percent FFS_L; $p=0.916$ over all rounds; $p>0.100$ for each round). Moreover, in both FFS_L and CAP_H we observe a (weakly) significant deviation from patient optimum even in the last round ($p\leq0.084$). Still, after round 1, in all conditions with competition, average deviations are significantly lower than in the respective conditions without competition ($p\leq0.030$ over all rounds; $p>0.100$ for at least 75 percent of the rounds per condition).

The findings on the competition conditions are confirmed when controlling for the dynamics of behavior in panel data models with random effects and clusters at the pair level. In these models, the adjusted deviation from the patient optimum is defined as dependent variable (see Table 2). To capture both effects of the experimental condition and dynamics over time, we consider as independent variables (i) the *round*, which represents the round of play; (ii) dummy variables with value 1 for conditions *CAP_H*, *FFS_L*, and *FFS_H (CAP_L*, *FFS_L*, and *FFS_H)*, respectively, while CAP_L (CAP_H) serves as baseline in Model 1 (Model 2); (iii) the interaction of condition and round to disentangle payment system and patient type specific effects of dynamics (*CAP_L x round, CAP_H x round*, *FFS_L x round*, *FFS_H x round*); and (iv) personal characteristics which are elicited by a questionnaire after the experiment. Personal characteristics incorporate a dummy variable for gender, which is 1 for *male* and 0 for female, and field of study with the dummy *econ* taking the value 1 for economics students and 0 for students of other fields. Both characteristics might be related to selfish behavior (e.g., Brosig-Koch et al., 2011, Carter and Irons, 1991, Croson and Gneezy, 2009).

Table 2: Panel data estimation of adjusted deviation from patient optimum

	Model 1	Model 2
Round	0.003	-0.033***
	(0.011)	(0.011)
CAP_L		-1.503***
		(0.463)
CAP_H	1.503***	
	(0.463)	
FFS_L	1.266***	-0.236
	(0.473)	(0.473)
FFS_H	0.243	-1.259***
	(0.463)	(0.462)
CAP_L x round		0.036**
		(0.016)
CAP_H x round	-0.036**	
	(0.016)	
FFS_L x round	0.063***	0.100***
	(0.017)	(0.017)
FFS_H x round	-0.016	0.020
	(0.016)	(0.016)
Male	0.012	0.012
	(0.091)	(0.091)
Econ	0.056	0.056
	(0.098)	(0.098)
Constant	-0.080	1.422***
	(0.332)	(0.330)
Groups, N	47, 1880	47, 1880
R-sq (overall)	0.18	0.18

Notes: Model 1 and Model 2 only differ in the specification of the baseline condition. Baseline for Model 1 is CAP_L and baseline for Model 2 is CAP_H. A negative coefficient reveals a reduction of deviation from patient-optimal treatment, thus, a better treatment quality. Cluster-robust standard errors in parentheses.

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

In line with previous results, model 1 shows that subjects deviate significantly more from patient-optimal treatment in conditions CAP_H and FFS_L than in the benchmark condition CAP_L. As expected from Figure 5, behavior in CAP_L and FFS_H does not differ significantly. Although payment systems are designed as mirror images of each other, we also find a general decreasing trend of deviations for CAP_H and an increasing trend for FFS_L, as the interaction terms of condition and round indicate. Model 2 further reveals that, initially, there is no significant difference between the two conditions. But, similar to model 1, the interaction term of condition and round in model 2 shows a significantly increasing trend of deviations for FFS_L compared to CAP_H. Apparently, subjects are more likely to coordinate on increasing the joint profit if incentives for this coordination are high and if this coordination implies providing *more* than what is optimal for the patient instead of providing *less* (for a graphical illustration, see Appendix C). Neither model 1 nor model 2 reveals a significant effect of personal characteristics like gender or field of study.

4.3 Tacit Collusion

We continue with analyzing behavior at a less aggregated level and focus on pairwise choices made in the competition conditions. We particularly examine the incidences of tacit collusion and distinguish between full collusion, coordination as well as attempts of full collusion, and attempts of coordination. *Full collusion* occurs if both subjects choose the joint profit-maximal treatment level q=10 (in FFS_L and FFS_H) or q=0 (in CAP_H and CAP_L). *Coordination* covers all pairwise equal deviation choices between the patient-optimal treatment level and the full collusion treatment level (i.e., both subjects deviate by 1, 2, 3, 4, 5, or 6 quantities, respectively, in FFS_L and CAP_H, and both deviate by 1 or 2 quantities, respectively, in FFS_H and CAP_L.). *Attempts of full collusion / coordination* relate to individual one-sided deviations from the patient-optimal treatment level. The observed frequencies of respective choices are summarized in Table 3. Overall, collusive behavior is rather rarely observed in our experiment. In each of the conditions CAP_L, CAP_H, and FFS_H full collusion occurs in less than three out of 240 cases. In FFS_L we observe full collusion in 21 cases, albeit concentrated in three pairs of subjects. In CAP_H and FFS_L a considerable number of subjects tries to fully collude or to coordinate but, particularly in CAP_H, often fails to succeed.

Table 3: Absolute frequency of full collusion, coordination, and attempts of collusion

Condition	# Rounds	# Pair Decisions	# Full Collusion	# Coordi-nation	# Attempts of full collusion	# Attempts of coordination
CAP_L	20	240	0 (0.0)	0 (0.0)	2 (0.4)	13 (2.7)
CAP_H	20	240	2 (0.8)	14 (5.8)	23 (4.8)	72 (15.0)
FFS_L	20	220	21 (9.5)	11 (5.0)	39 (8.9)	40 (9.1)
FFS_H	20	240	1 (0.4)	0 (0.0)	8 (1.7)	13 (2.7)
FS	15*	255	31 (12.2)	49 (19.2)	38 (7.45)	211 (41.4)

Note: Relative frequency in parentheses.

*Duration is not known to the subjects (they only knew that there would be a "large number" of rounds, p.117)

Table 3 also includes data obtained in a somewhat related Bertrand competition experiment run by Fouraker und Siegel (1963, FS). Bertrand competition is related to our setting as both are modeled as games of strategic complements. In their experiment, FS test whether repeated price competition between two sellers leads to competitive outcomes. To compare the incentives for tacit collision in FS's experimental set-up to those in our experiment, we apply the Friedman Index (see Friedman, 1971). This index is calculated as the difference between the profit in case of full collusion and the equilibrium profit (i.e., the potential gain from full collusion) divided by the difference between the maximum profit from unilateral defection and the profit from full collusion (i.e., the potential gain from defection). The higher the index is the higher is the monetary incentive to tacitly collude. We find that the index for FS (0.767) is higher than that for CAP_L and FFS_H (0.186), but lower than that for CAP_H and FFS_L (1.010). Irrespective of that, we observe a higher share of fully collusive outcomes, a higher share of coordinated outcomes, and a higher share of coordination attempts in FS than in any one of our four conditions. Possibly, tacitly collusive behavior in a medical service provision setting is not as frequent as in a conventional price competition setting. However, the comparison has to be interpreted with care as there are other differences in design that that might have affected the share of collusive decisions in FS (e.g., in their experiment subjects were not informed about the exact number of rounds to be played).

Finally, we turn to the determinants of treatment choices. Specifically, we investigate how individual characteristics and past experiences influence actual behavior. We regress a dummy for choosing the collusion quantity – which is one if the maximum (minimum) treatment level in FFS (CAP) is chosen – and a dummy for the patient-optimal treatment level on behavior in the previous round t-1, respectively. *L1.collusion* is an indicator variable that takes the value 1 if both subjects chose the collusion quantity in the previous round. The dummy

L1.attempt_collusion_j equals 1 if only the opponent chose the collusion quantity and potentially signaled his willingness to collude. *L1.patient_optimum* captures the patient-optimal treatment choice of both subjects in the previous round. The impact of the case that only the opponent chose the patient optimum is given by the regressor *L1.attempt_patient_optimum_j*. As the frequencies included in Table 3 suggest considerable differences between conditions, we also control for condition effects (*CAP_H*, *FFS_L*, *FFS_H*) and personal characteristics with the dummy variables *male* and *econ.*

Regression (1) in Table 4 shows the impact of previous behavior on the likelihood of choosing the collusion quantity. Full collusion in *t*-1 significantly increases the likelihood of maintaining the collusion quantity. However, a unilateral choice of the collusion quantity by the opponent does not significantly influence a subject's willingness to choose this level. Furthermore, mutual choices of the patient-optimal treatment level in the past round significantly reduce the likelihood of choosing the collusion quantity. Interestingly, in case of a unilateral choice of the patient optimum by the opponent, the likelihood of choosing the collusion quantity significantly increases. Possibly, subjects signaling to reach a collusive outcome tend to strengthen their signal in case they observed a patient-optimal choice by the opponent. The regression also reveals a significant influence of payment incentives: In conditions CAP_H and FFS_L (in which the long-run loss from a collapse of collusion is high compared to the short-term gain of defection) the likelihood of choosing the profit maximal treatment level significantly increases. Again, gender and field of study have no significant effect.

Table 4: Panel probit analysis on (1) collusion quantity and (2) patient-optimal treatment level, clustered at the pair level

	(1) Collusion quantity		(2) Patient-optimal treatment level	
L1.collusion	1.287***	(0.256)	-0.973***	(0.269)
L1.attempt_collusion_j	0.294	(0.211)	0.083	(0.174)
L1.patient_optimum	-0.496***	(0.177)	1.094***	(0.105)
L1.attempt_patient_optimum_j	0.386**	(0.163)	0.119	(0.118)
CAP_H	0.820**	(0.373)	-1.112***	(0.228)
FFS_L	1.380***	(0.370)	-1.165***	(0.233)
FFS_H	0.510	(0.384)	-0.156	(0.237)
Male	0.121	(0.152)	0.082	(0.100)
Econ	-0.160	(0.158)	-0.074	(0.108)
Constant	-2.641***	(0.380)	1.001***	(0.204)
Groups, N	47,1880		47,1880	
Rho	0.162[a]		0.151 [a]	
Log Likelihood	-282.608		-649.844	

[a] Significant at the 1% level using LR test.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.
Cluster-robust standard errors in parentheses.

Regression (2) analyzes the impact of previous behavior on the likelihood of choosing the patient-optimal treatment level. As expected and in line with previous results, this regression shows that full collusion in the previous round reduces the likelihood of treating patients optimally. Furthermore, mutual choices of the patient optimum increase the likelihood of choosing the patient-optimal treatment. In conditions in which incentives for collusion are high (CAP_H and FFS_L) the likelihood of choosing the patient-optimal treatment level significantly decreases.

5 Conclusion

In this study, we use a controlled laboratory experiment to investigate how competition in medical service provision affects patient outcomes. In line with the predictions of a model developed as a theoretical benchmark for this study, competition reduces the distortive impact of capitation and fee-for-service payment. Without competition we find underprovision with capitation payment and overprovision with fee-for-service payment, though less pronounced than predicted with pure profit maximization. Observed behavioral patterns are consistent with an average degree of physician altruism $\alpha \in [0,1]$. Thus, our results on conditions without competition are in line with previous experimental evidence on physician payment (e.g., Hennig-Schmidt et al., 2011, Godager and Wiesen, 2013, Brosig-Koch et al., 2015a,b).

With competition the deviations from patient-optimal treatment are lower than without competition. But even with finitely repeated competition, treatment choices still deviate from the patient optimum. Observed deviations depend on both, patient characteristics and physician payment. Deviations are higher for patients in need of a high level of medical treatment under capitation and for patients in need of a low level of medical treatment under fee-for-service. Moreover, deviations from patient-optimal treatment seem to be somewhat higher with fee-for-service than with capitation payment even though the two payment systems are symmetrically designed in our experiment.

Repeated competition seems to foster tacit collusion only when the long-run loss of a collapse of cooperation is high compared to the short-term gain of defection. Interestingly, deviations from patient-optimal treatment even increase with repetition under fee-for-service payment (while they decrease with repetition under capitation). Possibly, under competitive pressure subjects perceive deviating from the patient optimum by providing too many medical services (which increases individual profit under fee-for-service) less badly than by providing too little (which increases profits under capitation).

Nevertheless, the degree of tacit collusion observed in our study seems to be rather low compared to what is typically observed in price competition experiments. Apparently, medical service provision might be less prone to tacit collusion than decisions made in classical Bertrand environments. As such our controlled laboratory experiment provides support for the supposition that provider competition can have positive effects on patient health outcomes.

References

Allard, M., Jelovac, I., Léger, P. (2011). Treatment and Referral Decisions under Different Physician Payment Mechanisms. *Journal of Health Economics*, 30, 880-893.

Allard, M., Léger, P.T., Rochaix, L. (2009). Provider Competition in a Dynamic Setting. *Journal of Economics and Management Strategy*, 16, 457-486.

Biørn, E., Godager, G. (2010). Does quality influence choice of general practitioner? An analysis of matched doctor–patient panel data. *Economic Modelling*, 27(4), 842-853.

Brosig-Koch, J., Helbach, C., Ockenfels, A., Weimann, J. (2011). Still different after all these years: Solidarity behavior in East and West Germany. *Journal of Public Economics* 95, S. 1373-1376.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Wiesen, D. (2015a). Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. *Journal of Economic Behavior & Organization*, forthcoming.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Wiesen, D. (2015b). The Effects of Introducing Mixed Payment Systems for Physicians: Experimental Evidence. *Health Economics*, forthcoming.

Carter, J. R., Irons, M., 1991. Are economists different, and if so, why? *Journal of Economic Perspectives* 5 (2), 171–177.

Chalkley, M., Malcomson, J. M. (1998). Contracting for Health Services when Patient Demand does not Reflect Quality. *Journal of Health Economics*, 17, 1-19.

Choné, P., Ma, C. (2011). Optimal Health Care Contract under Physician Agency. *Annales d'Economie et de Statistique*, 101/202, 229-256.

Croson, R., Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature* 47 (2), 448-474.

Davidson, S., Manheim, L., Werner, S., Hohlen, M., Yudkowsky, B., Flemming, G. (1992). Prepayment with office-based physicians in publicly funded programs: results from children's medicaid program, *Pediatrics*, 89, 761–767.

Devlin, R. A., Sarma, S. (2008). Do physician remuneration schemes matter? The case of Canadian family physicians. *Journal of Health Economics*, 27(5), 1168-1181.

Dixon, P., Gravelle, H., Carr-Hill, R., J. Posnett, J. (1997). *Patient movements and patient choice: Report for National Health Service Executive*. London: York Health Economics Consortium.

Dranove, D., Jin, G. Z. (2010). Quality Disclosure and Certification: Theory and Practice, *Journal of Economic Literature* 48(4), 935–63.

Dulleck, U., Kerschbamer, R. (2006). On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods. *Journal of Economic Literature* 44(1), 5-24.

Dunn, A., Shapiro, A. H. (2015): Physician Competition and the Provision of Care: Evidence from Heart Attacks. *Federal Reserve Bank of San Francisco Working Paper* 2015-07.

Ellis, R.P. (1998). Creaming, skimping, dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics*, 17, 537-555.

Ellis, R. P., McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, 5(2), 129-151.

Ellis, R. P., McGuire, T. G. (1990). Optimal payment systems for health services. *Journal of Health Economics*, 9(4), 375-396.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science* 326(5952), 535-538.

Fischbacher, U. (2007). Z-tree: Zurich Toolbox for Readymade Economic Experiments, *Experimental Economics*, 10, 171-178.

Fouraker, L.A., Siegel, S. (1963). *Bargaining Behavior*. McGraw-Hill, New York.

Fréchette, G. R., Schotter, A. (Eds.). (2015). *Handbook of Experimental Economic Methodology*. Oxford University Press.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames, *Review of Economic Studies*, 1-12.

Gaynor, M., Gertler, P. (1995). Moral hazard and risk spreading in partnerships. *The RAND Journal of Economics*, 591-613.

Gaynor, M., Town, R. J. (2012). Competition in Health Care Markets. In: *Handbook of Health Economics* Vol. 2, ed. by Pauly, McGuire, and Barros, North-Holland, Amsterdam (The Netherlands), 499-637.

Godager, G., Wiesen, D. (2013). Profit or Patients' Health Benefit? Exploring the Heterogeneity in Physician Altruism. *Journal of Health Economics*, 32, 1105-1116.

Godager, G., Iversen, T., Ma, C. T. A. (2015). Competition, gatekeeping, and health care access. *Journal of Health Economics*, 39, 159-170.

Gosden, T., Forland, F., Kristiansen, I. S., Sutton, M., Leese, B., Giuffrida, A., Sergison, M., Pedersen, L. (2001). Impact of payment method on behaviour of primary care physicians: a systematic review. *Journal of Health Services Research & Policy*, 6(1), 44-55.

Green, E. P. (2014). Payment systems in the healthcare industry: An experimental study of physician incentives. *Journal of Economic Behavior & Organization*, 106, 367-378.

Gravelle, H., Masiero, G. (2000). Quality incentives in a regulated market with imperfect information and switching costs: capitation in general practice. *Journal of Health Economics*, 19, 1067-1088.

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. in V. M. Kurt Kremer (ed.), *Forschung und Wissenschaftliches Rechnen 2003*. GWDG Bericht 63, Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, 79-93.

Grytten, J., Sørensen, R. (2001). Type of contract and supplier-induced demand for primary physicians in Norway. *Journal of Health Economics*, 20(3), 379-393.

Hennig-Schmidt, H., Selten, R., Wiesen, D. (2011). How payment systems affect physicians' provision behavior – an experimental investigation. *Journal of Health Economics*, 30(4), 637-646.

Herbst, D., Mas, A. (2015). Peer effects on worker output in the laboratory generalize to the field. *Science* 350(6260), 545-549.

Huck, S., Lünser, G. K., Spitzer, F., Tyran, J. R. (2014). Medical Insurance and Free Choice of Physician Shape Patient Overtreatment. A Laboratory Experiment. *Discussion Paper No. 14-19*, Department of Economics, University of Copenhagen.

Hurley, J., Labelle, R. (1995). Relative fees and the utilization of physicians' services in Canada. *Health Economics*, 4(6), 419-438.

Hutchison, B., Birch, S., Hurley, J., Lomas, J., Stratford-Devai, F. (1996). Do physician-payment mechanisms affect hospital utilization? A study of Health Service Organizations in Ontario. CMAJ: *Canadian Medical Association Journal*, 154(5), 653.

Iversen, T., Lurås, H. (2006). *Capitation and incentives in primary care*. The Elgar Companion to Health Economics, 269.

Iversen, T., Ma, C. T. A. (2011). Market conditions and general practitioners' referrals. *International Journal of Health Care Finance and Economics*, 11(4), 245-265.

Karlsson, M. (2007). Quality incentives for GPs in a regulated market. *Journal of Health Economics*, 26, 699-720.

Levitt, S. D., List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic Perspectives*, 153-174.

Ma, C. (1994): Health Care Payment Systems: Cost and Quality Incentives, *Journal of Economics and Management Strategy*, 3, 93-112.

McGuire, T. G. (2000). Physician Agency. In: *Handbook of Health Economics*, Vol. 1 A, ed. by Cuyler and Newhouse, North-Holland, Amsterdam (The Netherlands), 461-536.

Potters, J., Suetens, S. (2013). Oligopoly experiments in the current millennium. *Journal of Economic Surveys*, 27(3), 439-460.

Pike, C. (2010). An empirical analysis of the effects of GP competition. *MPRA Paper No. 27613*.

Salisbury, C. J. (1989). How do people choose their doctor?. *British Medical Journal*, 299(6699), 608.

Appendix A: Instructions + Comprehension Questions (Competition Conditions)

Welcome to the Experiment!

Preface

You are participating in an economic experiment on decision behavior. You and the other participants will be asked to make decisions for which you can earn money. Your payoff depends on both your decisions and the decisions of the other participants. At the end of the experiment, your payoff will be converted to Euro and paid to you in cash. During the experiment, all amounts are presented in the experimental currency Taler. 100 Taler equals 5 Euro. The experiment will take about 90 minutes. All participants receive the same instructions.

Please read the following instructions carefully. We will approach you in about five minutes to answer any questions you may have. If you have questions at any time during the experiment, please raise your hand and we will come to you.

Decision Situations

In each round you take on the role of a physician and decide on medical treatment for a patient. The total number of patients which can receive medical treatment you will find out in section “patients”. At the beginning of the experiment you will be randomly matched with another participant who will also take on the role of a physician and decide on medical treatment for patients. The experiment will consist of 20 decision rounds. During the experiment you solely interact with the same participant.

In each round you determine the quantity of medical treatment for each patient. That is, all patients in this round will be treated with the identical quantity determined by you. Your decision is to provide each patient with a quantity of 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 medical services. Every quantity of medical service yields a particular benefit for the patient. The benefit resulting from a specific quantity of medical services is identical for you and the other physician.

Patients

In each of the 20 rounds four patients can get medical treatment. The following applies to each of the 20 rounds.

Two out of four patients are regular patients, whereas one regular patient is assigned to you and the one is assigned to the other physician. Regular patients always remain with the physician to whom they were initially assigned to, independently of the number of medical services you and the other physician provide. The other two patients are patients who are undecided. That is, they have not yet been set to a treating physician. The following applies to the undecided patients.

- They get the treatment from you if the medical treatment provided by you leads to a higher benefit than the medical treatment of the other physician.
- They get the treatment from the other physician if his medical treatment leads to a higher benefit than your treatment.
- If both patients receive the same benefit, they will split equally between both physicians.

You and the other physician independently decide on the number of medical services for all patients. In particular, the number of medical services you provide applies to all of your patients. The patients who had been undecided so far will then be assigned to a physician according the benefit they receive.

<u>Profit</u>

In each round you receive a fee-for-service (capitation) remuneration for treating the patients. Your remuneration increases with the amount of medical treatment (is irrespective of the amount of medical treatment) you provide. You also incur costs for treating the patients, which likewise depend on the quantity of services you provide. Your profit per patient treated is calculated by subtracting these costs from the fee-for-service (capitation) remuneration. Your total profit for each round is then the profit per patient multiplied with the number of patients you have treated.

Every quantity of medical service yields a particular benefit for the patient. Hence, in choosing the medical services you provide, you determine not only your own profit but also the patient's benefit.

In each round you will receive detailed information on your screen (see below) about the number of regular patients and the number of patients which are undecided. You also receive information on the amount of your fee-for-service (capitation) remuneration per patient and – for each possible amount of medical treatment – your costs, profit as well as the benefit for the patients.

After each round you will receive information on your screen (see above) about your decision, the number of medical services provided by the other physician, as well as the resulting number of patients treated by each physician. Furthermore, this information will be displayed for all previous rounds.

Payment

At the end of the experiment your total profit out of each round will be summed up and paid to you in cash.

For this experiment, no patients are physically present in the laboratory. Yet, the patient benefit of the four patients in each of the 20 rounds does accrue to real patients: The added patient benefit resulting from the medical treatment of the four patients in each of the 20 rounds will be transferred to the Christoffel-Blindenmission Deutschland e.V., 64625 Bensheim, an organization which funds the treatment of patients with eye cataract.

The transfer of money to the Christoffel-Blindenmission Deutschland e.V. will be carried out after the experiment by the experimenter and one participant. The participant completes a money transfer form, filling in the total patient benefit (in Euro) resulting from the decisions made by all participants. This form prompts the payment of the designated amount to the Christoffel-Blindenmission Deutschland e.V. by the University of Duisburg-Essen’s finance department. The form is then sealed in a postpaid envelope and posted in the nearest mailbox by the participant and the experimenter.

After the entire experiment is completed, one participant is chosen at random to oversee the money transfer to the Christoffel-Blindenmission Deutschland e.V. The participant receives an additional compensation of 5 Euro for this task. The participant certifies that the process has been completed as described here by signing a statement which can be inspected by all participants at the office of the Chair of Quantitative Economic Policy. A receipt of the bank transfer to the Christoffel-Blindenmission Deutschland e.V. may also be viewed here.

Comprehension Questions

Prior to the decision rounds we kindly ask you to answer a few comprehension questions. They are intended to help you familiarize yourself with the decision situations. If you have any questions about this, please raise your hand. The experiment will begin once all participants have answered the comprehension questions correctly.

Comprehension Questions: CAP_L (FFS_L)

Number of your regular patients: 1
Number of regular patients of the other physician: 1
Number of undecided patients: 2

Quantity of medical treatment per patient	Capitation (Fee-for-service) per patient (in Taler)	Costs per patient (in Taler)	Profit per patient (in Taler)	Benefit of the patient (in Taler)
0	20.00 (0.00)	0.00	20.00 (0.00)	14.00
1	20.00 (4.00)	0.20	19.80 (3.80)	16.00
2	20.00 (8.00)	0.80	19.20 (7.20)	18.00
3	20.00 (12.00)	1.80	18.20 (10.20)	20.00
4	20.00 (16.00)	3.20	16.80 (12.80)	18.00
5	20.00 (20.00)	5.00	15.00 (15.00)	16.00
6	20.00 (24.00)	7.20	12.80 (16.80)	14.00
7	20.00 (28.00)	9.80	10.20 (18.20)	12.00
8	20.00 (32.00)	12.80	7.20 (19.20)	10.00
9	20.00 (36.00)	16.20	3.80 (19.80)	8.00
10	20.00 (40.00)	20.00	0.00 (20.00)	6.00

1. Assume that a physician wants to provide 9 quantities of medical treatment for the patients depicted above.

1 a) What is the capitation (fee-for-service) per patient?

1 b) What are the costs per patient?

1 c) What is the profit per patient?

1 d) What is the patient benefit per patient?

2. Assume that a physician wants to provide 9 quantities of medical treatment for the patients depicted above. The other physician wants to provide 2 quantities of medical treatment for these patients.

2 a) How many regular patients would you treat?

2 b) How many undecided patients would you treat?

2 c) How many patients would you treat in total?

2 d) What is your total profit?

Comprehension Questions: CAP_H (FFS_H)

Number of your regular patients: 1
Number of regular patients of the other physician: 1
Number of undecided patients: 2

Quantity of medical treatment per patient	Capitation (Fee-for-service) per patient (in Taler)	Costs per patient (in Taler)	Profit per patient (in Taler)	Benefit of the patient (in Taler)
0	20.00 (0.00)	0.00	20.00 (0.00)	6.00
1	20.00 (4.00)	0.20	19.80 (3.80)	8.00
2	20.00 (8.00)	0.80	19.20 (7.20)	10.00
3	20.00 (12.00)	1.80	18.20 (10.20)	12.00
4	20.00 (16.00)	3.20	16.80 (12.80)	14.00
5	20.00 (20.00)	5.00	15.00 (15.00)	16.00
6	20.00 (24.00)	7.20	12.80 (16.80)	18.00
7	20.00 (28.00)	9.80	10.20 (18.20)	20.00
8	20.00 (32.00)	12.80	7.20 (19.20)	18.00
9	20.00 (36.00)	16.20	3.80 (19.80)	16.00
10	20.00 (40.00)	20.00	0.00 (20.00)	14.00

1. Assume that a physician wants to provide 2 quantities of medical treatment for the patients depicted above.

1 a) What is the capitation (fee-for-service) per patient?

1 b) What are the costs per patient?

1 c) What is the profit per patient?

1 d) What is the patient benefit per patient?

2. Assume that a physician wants to provide 2 quantities of medical treatment for the patients depicted above. The other physician wants to provide 9 quantities of medical treatment for these patients.

2 a) How many regular patients would you treat?

2 b) How many undecided patients would you treat?

2 c) How many patients would you treat in total?

2 d) What is your total profit?

Appendix B: Proof of the Proposition

By symmetry of the experimental designs FFS_L and CAP_H and of FFS_H and CAP_L, we may w.l.o.g. restrict attention to the experimental conditions FFS_L and FFS_H. It suffices to show that, for any degree of altruism $\alpha_1 \in [0,1]$, physician 1 maximizes utility by choosing q^*, given that physician 2 picks the equilibrium level of treatment q^*. To this end, let $q_1 \neq q^*$ denote an alternative treatment level of physician 1. Then, a similar argument applies to physician 2.

(Existence) Consider the experimental condition FFS_L, i.e. we have $q^* = q^L = 3$. Define

$$\Delta(q_1, \alpha_1) := U_1(q^*, q^*; \alpha_1) - U_1(q_1, q^*; \alpha_1).$$

First, if $\alpha_1 = 0$ then $U_1(q^*, q^*; 0) = 2\pi^{FSS}(3) = 10.2$ and $U_1(q_1, q^*; 0) = \pi(q_1) \leq \pi(10) = 10$ and hence $\Delta(q_1, 0) > 0$ for all $q_1 \neq q^*$. Second, if $\alpha_1 = 1$ then $U_1(q^*, q^*; 1) = 2B(q^*) = 20$ and $U_1(q_1, q^*; 1) = B(q_1) \leq 9$ and hence $\Delta(q_1, 1) > 0$ for all $q_1 \neq q^*$. Finally, for any $\alpha_1 \in (0,1)$, we obtain $U_1(q^*, q^*; \alpha_1) \geq 10.2$ and $\max_{q_1 \neq q^*} U_1(q_1, q^*; \alpha_1) < 10$. It therefore follows that $\Delta(q_1, \alpha_1) > 0$ for all $q_1 \neq q^*$ and all $\alpha_1 \in (0,1)$. Thus, (q^*, q^*) represents a strict Nash equilibrium in FFS_L for all $(\alpha_1, \alpha_2) \in [0,1]^2$.

A similar argument shows that $(q^*, q^*) = (7,7)$ constitutes a Nash equilibrium in FFS_H.

(Uniqueness) It remains to be shown that (q^*, q^*) represents a *unique* Nash equilibrium of the stage game. Without loss of generality, we can restrict attention to experimental condition FFS. We show (1) that no other symmetric Nash equilibrium exists and (2) that no asymmetric equilibrium exists either.

Ad (1): Let (q, q) represent a Nash equilibrium such that $q \neq q^*$. If $q < q^*$ then $q_1 = q + 1$ entails $B(q+1) > B(q)$ and hence $n(q+1, q) > n(q, q)$. Therefore, $\pi^{FFS}(q+1) > \pi^{FFS}(q)$ implies $U_1(q+1, q; \alpha_1) > U_1(q, q; \alpha_1)$ for any $\alpha_1 \in [0,1]$. Thus, (q, q) does not represent a Nash equilibrium for $q < q^*$.

On the other hand, if $q > q^*$ then $q_1 = q - 1$ entails $B(q-1) > B(q)$, which implies that (q, q) does not represent an equilibrium for $\alpha_1 = 1$. Moreover, it follows that $n(q-1, q) = 3 > n(q, q) = 2$, but also $\pi^{FFS}(q-1) < \pi^{FFS}(q)$. Notice, however, that $\pi_1(q-1, q) = 3\pi^{FFS}(q-1) > \pi_1(q, q) = 2\pi^{FSS}(q)$. To show this, we set

$$\Delta(q) := 3\pi^{FFS}(q-1) - 2\pi^{FSS}(q) = \frac{1}{10}(-q^2 + 26q - 63).$$

Since $\Delta(q)$ is strictly increasing on Q, it follows from $q > q^* \geq 3$ that $\Delta(q) > \Delta(3) = 6/10 > 0$, i.e. we have $\pi_1(q-1,q) > \pi_1(q,q)$ and thus (q,q) does not represent an equilibrium for $\alpha_1 = 0$. Combined with $B(q-1) > B(q)$ this implies that $U_1(q-1,q;\alpha_1) > U_1(q,q;\alpha_1)$ for any $\alpha_1 \in (0,1)$, since physician 1's utility is linear in α_1. Thus, (q,q) does not represent a Nash equilibrium for $q > q^*$ either.

Ad (2): Let (q_1,q_2) represent a Nash equilibrium such that $q_1 \neq q_2$. Observe that $q_1 = q^*$ or $q_2 = q^*$ cannot be part of an asymmetric equilibrium since (q^*,q^*) represents a *strict* Nash equilibrium. Hence suppose that $q_1 \neq q^*$ and $q_2 \neq q^*$. Without loss of generality, let $q_1 < q_2$. If $q_1 < q^*$ then physician 1 can increase utility by choosing q^* instead. To see this, notice that $B(q^*) > B(q_1)$, $n(q^*,q_2) \geq n(q_1,q_2)$, and $\pi^{FFS}(q^*) > \pi^{FFS}(q_1)$. It thus follows that $U_1(q^*,q_2;\alpha_1) > U_1(q_1,q_2;\alpha_1)$ for any $\alpha_1 \in [0,1]$ in contradiction to (q_1,q_2) representing a Nash equilibrium.

On the other hand, if $q_1 > q^*$, then physician 2 can increase utility by choosing $q_1 - 1$ instead. To see this, observe first that patient benefit strictly increases, $B(q_1 - 1) > B(q_2)$. Hence, (q,q) does not represent an equilibrium for $\alpha_2 = 1$. Secondly, this implies $n(q_1, q_1 - 1) = 1$ whereas $n(q_1,q_2) = 3$, i.e. physician 2 attracts two additional patients. What is more, exploiting the strictly positive monotonicity of $\pi^{FSS}(q)$ reveals this deviation to be profitable:

$$\begin{aligned}\pi_2(q_1,q_1-1) &= 3\pi^{FSS}(q_1-1) \geq 3\pi^{FSS}(q^*) \geq 3\pi^{FSS}(3) = 15.3 \\ &> 10 = \pi^{FSS}(10) \geq \pi^{FSS}(q_2) = \pi_2(q_1,q_2).\end{aligned}$$

Therefore, (q,q) does not represent a Nash equilibrium for $\alpha_2 = 0$. Since physician 2's utility is linear in α_2, it hence follows that $U_2(q_1,q_1-1;\alpha_2) > U_2(q_1,q_2;\alpha_2)$ for any $\alpha_2 \in (0,1)$. Thus, (q_1,q_2) cannot represent a Nash equilibrium for $q_1 > q^*$ either. □

Appendix C: Frequency of Deviations from Patient-optimal Treatment per Round