RUHR

ECONOMIC PAPERS

Jeannette Brosig-Koch
Heike Hennig-Schmidt
Nadja Kairies-Schwarz
Daniel Wiesen

The Effects of Introducing Mixed Payment Systems for Physicians – Experimental Evidence

UNIVERSITÄT DUISBURG ESSEN

#543

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

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #543

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-621-5

Ruhr Economic Papers #543

Jeannette Brosig-Koch, Heike Hennig-Schmidt,
Nadja Kairies-Schwarz, and Daniel Wiesen

The Effects of Introducing Mixed Payment Systems for Physicians – Experimental Evidence

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.

http://dx.doi.org/10.4419/86788621
ISSN 1864-4872 (online)
ISBN 978-3-86788-621-5

Jeannette Brosig-Koch, Heike Hennig-Schmidt, Nadja Kairies-Schwarz, and Daniel Wiesen[1]

The Effects of Introducing Mixed Payment Systems for Physicians – Experimental Evidence

Abstract

Mixed payment systems have become a prominent alternative to paying physicians through fee-for-service and capitation. While theory shows mixed payment systems to be superior, empirically, causal effects on physicians' behavior are not well understood when introducing mixed systems. We systematically analyze the influence of fee-for-service, capitation, and mixed payment systems on physicians' service provision. In a controlled laboratory setting, we implement an exogenous variation of the payment method. Participants, in the role of physicians, in the lab (N=213) choose quantities of medical services affecting patients' health outside the lab. Behavioral data reveal significant overprovision of medical services under fee-for-service and significant underprovision under capitation, though less than predicted when assuming profit-maximization. Introducing mixed payment systems significantly reduces deviations from patient-optimal treatment. Responses to incentive systems can be explained by a behavioral model capturing physician altruism. We find substantial heterogeneity in physician altruism. Our results hold for medical and non-medical students.

JEL Classification: C91, I11

Keywords: Fee-for-service; capitation; mixed payment systems; physician altruism; laboratory experiment

February 2015

1 Jeannette Brosig-Koch, UDE and CINCH; Heike Hennig-Schmidt, University Bonn and University Oslo; Nadja Kairies-Schwarz, UDE and CINCH; Daniel Wiesen, University Köln and University Oslo. *– We are grateful for valuable comments and suggestions by Steen Altmann, Pedro Barros, Randy Ellis, Dirk Engelmann, Geir Godager, Laszlo Goerke, Oliver Gürtler, Jerry Hurley, Tor Iversen, Sara Machado, Banu Olcay, Dirk Sliwka, Reinhard Selten, Becky Staiger, and conference and seminar participants at ECHE Zurich 2012, ESA New York 2012, ESA European Conference Cologne 2012, dggö Essen 2013, Symposium of the German Economic Association of Business Administration in Magdeburg 2013, U Oslo, U Trier, 1st Workshop on Behavioral and Experimental Health Economics in Oslo 2014, and ASHEcon LA 2014. We thank Emanuel Castillo for his programming assistance, Andrew Kinder for his research assistance, and Johanna Kokot as well as Christoph Helbach for helping to conduct the experiments. Financial support by DFG (German Research Foundation, grant: BR 2346/2-1/2) is gratefully acknowledged. This paper is a completely revised version of the discussion paper. "How to Improve Patient Care? An Analysis of Capitation, Fee-for-Service, and Mixed Payment Schemes for Physicians" (Ruhr Economic Papers, REP #412).* *– All correspondence to: Daniel Wiesen, Universität Köln, Albertus-Magnus-Platz, 50923 Köln, Germany, e-mail: daniel.wiesen@uni-koeln.de*

1 Introduction

Understanding how physicians respond to changes in the payment method is important for policy makers and researchers alike. Establishing the causal impact of a payment system variation in the field is difficult, however, due to the likely endogeneity of institutions or self-selection problems (see, e.g., Baicker and Goldman, 2011). This paper contributes to a better understanding of how payment systems affect physicians' behavior. To this end, we use controlled laboratory experiments in the spirit of the physician decision-making setting by Hennig-Schmidt et al. (2011) as a complementary approach to field studies.[1] We introduce ceteris-paribus variations in our experiment by manipulating the payment system while keeping all other variables (e.g., patient characteristics) constant. Thus, we are able to analyze the causal effect of a payment system variation on physicians' behavior.

At one extreme, we implement a non-blended fee-for-service (FFS) system, which is still the most common form of paying physicians in many health care markets. In FFS, physicians receive a fee for each service they provide and are thus incentivized to provide too many services (overprovision). At the other end of the spectrum, we consider capitation (CAP), which pays physicians a lump-sum for each enrolled patient, embedding an incentive to provide too few services (underprovision). To curtail health care spending in the field, mixed payment systems have been introduced that comprise both FFS and CAP components. Managed care, for example, relies heavily on supply-side incentives to control costs by changing physician payment methods (see, e.g., Ellis and McGuire, 1993). To mimic the introduction of mixed payment systems at a within-subject level, we systematically change physicians' remuneration from either FFS or CAP to different mixed systems which vary the respective weighting put on fee-for-service and lump-sum components. As in some mixed systems, the fee is set below marginal cost, behavioral responses to these systems can be interpreted as consequences of the degree of supply-side cost sharing.[2]

All participants in our experiment decide in the role of 'physicians' on the quantity of medical services for different 'patients'. Henceforth we use these labels to indicate the roles

[1] Applying the experimental method in health economics is a rather new approach. It is still in its infancy, even though Fuchs (2000) and Frank (2007) more than a decade ago proposed incorporating behavioral and experimental methods to complement traditional approaches in health economics research. Recently, a growing number of research topics in health economics have been addressed by laboratory experiments, indicating its increasing importance in this field. Laboratory experiments analyze, for example, health insurance choices (Schram and Sonnemans, 2011), health care financing (e.g., Buckley et al., 2012), the allocation of medical resources (e.g., Ahlert et al., 2012), the salience of the Hippocratic Oath (Kesternich et al., 2014), and other-regarding behavior and motivations (Hennig-Schmidt and Wiesen, 2014).

[2] Supply-side cost sharing has become a prominent option to control health care costs (Ellis and McGuire, 1993, p.135), as opposed to demand-side cost sharing, where patients make co-payments or pay deductibles.

in our experiment. A physician's quantity choice determines the physician's own profit and a patient's health benefit. Decisions are incentivized by monetary rewards determined by the respective payment method. Real patients' health outside the lab is affected by these decisions. Further, participants are randomly assigned to the experimental conditions, thereby excluding selection biases. We also control for subjects' medical background as both medical and non-medical students participate in our experiment.

Our goal of establishing the effects a variation from FFS or CAP to mixed payment systems has on physicians' behavior requires that payment systems are designed to ensure comparability. To this end, payment parameters are chosen such that maximum profits and (for Pareto-efficient choices) marginal profits are equivalent across payment systems. Moreover, patient health benefits are concave, with a global optimum on the physicians' quantity choice range. Thus, we are able to identify over- and underprovision of medical services in our behavioral data. The patient benefit functions are constructed in a completely symmetric way such that equal marginal effects of underprovision and overprovision are identical. This implies that tradeoffs between profits and patient health benefits are comparable across payment systems. A physician paid by FFS treating a patient with a high severity of illness faces equivalent tradeoffs, for example, compared to a physician paid by CAP treating a low-severity patient. These novel design features differentiate our study from earlier experiments (e.g., Hennig-Schmidt et al., 2011). Further, as the patient population is kept constant across payment systems, we are able to investigate—within the confines of our experiment—the impact of a payment system variation on patients' health benefits, a relevant topic in the recent empirical literature (e.g., Clemens and Gottlieb, 2014). Finally, the experimental design also allows us to analyze to what extent observed treatment decisions can be accounted for by physician altruism, i.e., the weight a physician attaches to patient's health benefit.[3]

In their seminal model, in which a physician values both her own profit and the patient's health benefit, Ellis and McGuire (1986) show that mixed payment systems induce the optimal level of health care services. We use a variant of their model to derive behavioral predictions for our experiment. According to the predictions based on our experimental parameters, introducing mixed payment systems should reduce the deviations from the patient-optimal provision level. Moreover, underprovision and overprovision should be less pronounced with a higher degree of physician altruism.

Our behavioral data confirm the theoretical predictions. We find less overprovision

[3] In the theoretical health economics literature, the weight the utility-maximizing physician attaches to the patient's health benefit is often interpreted as physician altruism (see, e.g., Ellis and McGuire, 1986; McGuire, 2000). Since Arrow (1963) highlighted the importance of the physician's benevolent motive to care for a patient when describing physician behavior, the altruistic-physician assumption has become quite common in modeling physicians' behavior (see, e.g., Ellis and McGuire, 1986, 1990; Chalkley and Malcomson, 1998; Choné and Ma, 2011).

in mixed payment systems compared to FFS and less underprovision in mixed payment systems compared to CAP. Moreover, a higher (lower) weight on the lump-sum component in mixed payment systems yields a further reduction in overprovision (underprovision) of medical services. As a result, patients' health benefits are significantly higher in mixed payment systems. In line with the assumption of altruistic preferences, observed provision behavior varies with patients' characteristics. However, the degree of altruism differs substantially among participants. This finding complements recent studies reporting heterogeneity in physician altruism (e.g., Godager and Wiesen, 2013).

The remainder of the paper is organized as follows. Section 2 refers to the related literature. We present our experimental design and procedure in Section 3. In Section 4, we derive behavioral predictions for our experiment. Section 5 presents our results, and Section 6 concludes.

2 Related literature

Mixed payment systems, in which physicians receive a (reduced) fee for each provided service and a lump-sum payment for each enrolled patient, are hypothesized to mitigate non-optimal service provision. Ellis and McGuire (1986) show that mixed payment systems can be designed such that the optimal level of health care services is induced. That means, physicians paid by FFS are expected to reduce the quantity of medical services, and, thus decrease overprovision when shifting to a mixed payment (e.g., Ellis and McGuire, 1986, 1990). In contrast, compared to CAP a mixed payment system should induce physicians to increase medical services and, thus, reduce the incentive to provide too few medical services. We will use a variant of Ellis and McGuire's (1986) model to derive behavioral hypotheses for the parameters of our experiment.

Empirically, however, it is still not well understood how physicians respond to mixed payment systems compared to non-blended FFS and CAP systems and how patients' health is affected (see, e.g., Kantarevic et al., 2011). A few studies indicate that, compared to FFS, mixed payment systems lead to reduced quantity and an increased quality of medical services (e.g., Dumont et al., 2008; Kralj and Kantarevic, 2013). Comparing physicians' behavior in CAP and in a mixed system, Krasnik et al. (1990) report that physicians expand their services after a fee-for-service component has been included in a CAP system. The evidence from the field suggests positive consequences of introducing mixed payment systems. Yet the observational data used to analyze these effects entail a lack of control, as existing payment systems are typically adopted endogenously, which renders causal inferences from a change of the payment method on physician behavior rather difficult (see e.g., Falk and Heckman, 2009). Using the experimental economics method allows us to vary the payment system exogenously at a within-subject level, thus mimicking the change from a non-blended to a mixed payment system.

Some empirical studies demonstrate that physicians respond to enhanced fees in an FFS system with an increase in the volume of services (e.g., Yip, 1998; Kantarevic et al., 2011). Clemens and Gottlieb (2014) find that physicians respond to fee changes mostly for elective services. Not so much is known, empirically, about the behavioral responses to varying degrees of supply-side cost sharing. Our systematic variation of payment systems also provides important insights into the effects of supply-side cost sharing, which is a popular means of responding to rising health care costs (see, e.g., Ellis and McGuire, 1993; Baicker and Goldman, 2011).

Our study also relates to a stream of literature investigating physician altruism. In his influential paper, Arrow (1963) coined the importance of the physician's other-regarding motive to care for a patient when describing physician behavior. In subsequent research, several theoretical papers model the physician as deriving utility from both her own profit and the patient's health benefit (see, e.g., Ellis and McGuire, 1990; Chalkley and Malcomson, 1998). The weight the physician attaches to the latter is often interpreted as physician altruism (see, e.g., McGuire, 2000). The role of physician altruism is particularly emphasized in more recent theoretical papers. For example, Jack (2005) derives an optimal menu of cost-sharing schemes for physicians with unknown altruism. Siciliani (2009) models the impact of performance pay on the provision of medical services when providers differ in altruism. In Choné and Ma (2011), the optimal payment mechanism depends on physician altruism. Treatment and referral decisions of altruistic physicians under a gatekeeping regime are investigated by Allard et al. (2011). Liu and Ma (2013) study delegation of treatment plans.

Empirical research investigating physician altruism is scarce. Only a few empirical studies with physicians and medical students, mainly discrete-choice experiments, trace physicians' financial motivation in an indirect way by using proxy variables (see, e.g., Rizzo and Zeckhauser, 2007; Hanson and Jack, 2010). Scott and Sivey (2013) examine how physicians' characteristics are associated with monetary motivation. Common to these studies is that tradeoffs between physicians' income and patients' health benefit are not explicitly considered. The only study that analyzes tradeoffs between profit and patient health benefit is Godager and Wiesen (2013), who explore the heterogeneity in physician altruism using Hennig-Schmidt et al.'s (2011) experimental data. They find substantial heterogeneity in physician altruism using mixed and multinomial logit regressions. Our study differs from the latter in that we analytically infer physician altruism from subjects' choices in the experiment, employing Ellis and McGuire's (1986) model framework.

In sum, we contribute to the literature in three ways. First, our experimental design allows us to collect behavioral data based on the exogenous introduction of mixed payment systems. Second, we systematically test behavioral predictions from Ellis and McGuire's (1986) seminal model—and to the best of our knowledge we are the first to do so—by comparing non-blended and mixed payment systems. Finally, we provide further evidence on the heterogeneity of physician altruism.

3 Experimental design

3.1 Basic setup and decision situation

In our experiment, all subjects decide in the role of physicians on the provision of medical services. We randomly assign subjects to different payment conditions. In the experimental conditions, subjects participate in two subsequent payment systems. First, they are incentivized by non-blended FFS or CAP. Second, subjects are paid by a mixed system comprising both FFS and CAP-components. We thus exogenously vary the payment method at a within-subject level.

In all payment systems, physician i decides on the quantity of medical services $q \in [0, 10]$ for nine different patients ($j = 1, \ldots, 9$). Patients differ in illnesses $k \in \{A, B, C\}$ and in severities of illness $l \in \{x, y, z\}$. Patients are assumed to be passive and fully insured, accepting each level of medical service provided by the physician. Patients are the same in all payment conditions. The patient population for which a physician chooses services thus remains constant.

Physician i's payment is $R(q) = \mu L + (1 - \mu)pq$, with L being a lump-sum payment per patient, p a fee per service rendered to a patient and $\mu \in [0, 1]$ the weight on the lump-sum component, which is often interpreted as the degree of supply-side cost sharing. In the experiment, μ, L, and p are varied systematically (see Subsection 3.2). For example, in non-blended FFS and CAP, μ is 0 and 1, respectively. Physician i's profit is

$$\pi(q) = \mu L + (1 - \mu)pq - c(q), \tag{1}$$

with $L, p > 0$, $c'(q) > 0$ and $c''(q) > 0$. In the experiment, $c(q) = q^2/10$ for all payment systems.

When deciding on q, physician i simultaneously determines her own profit $\pi(q)$ and the health benefit $B(q)$ of patient j. Common to all nine patient health benefits is a global optimum at q^* on $q \in (0, 10)$. More formally, the patient health benefit employed in our experiment is

$$B(q) = \begin{cases} B_0 + \theta q & \text{if } q \leq q^* \\ B_1 - \theta q & \text{if } q \geq q^*, \end{cases} \tag{2}$$

with $B_0, B_1 \geq 0$ and $\theta > 0$. The patient's health benefit is varied systematically for the patients' illness k and severity of illness l. In particular, for illnesses A and B $\theta = 1$

and for illness C $\theta = 2$. For illnesses A, B, C the maximum health benefit is $B_{Al}(q^*) = 7$, $B_{Bl}(q^*) = 10$, and $B_{Cl}(q^*) = 14$, respectively. The patient-optimal quantity q^* varies with severities of illness l. For low (x), intermediate (y), and severe (z) severities, the patient-optimal quantities are $q^* = 3$, $q^* = 5$, and $q^* = 7$, respectively. Figure 2 in Appendix A.2 illustrates patient health benefits.[4] Knowing the patient-optimal quantity q^* allows us to analyze overprovision and underprovision of medical services. Moreover, the symmetric design of patient health benefits implies that the marginal effects (i.e., absolute value of $B'(q)$) of overprovision and underprovision are equivalent.[5]

All parameters of the experiment are common knowledge. In particular, when making their quantity choices physicians are aware of costs, payment, profit, and the patient's health benefits for each quantity (for an illustration of the decision situation, see the instructions in Appendix A.1). Therefore, behavioral patterns like defensive medicine (see, e.g., Kessler and McClellan, 1996) can be neglected, as the impact of a quantity choice on the patients' health benefit is known to the experimental subjects.

While all participants in the experiment make decisions in the role of physicians for abstract patients in the lab, real patients' health outside the lab is affected by their choices. Subjects are informed that the monetary equivalent of the patient health benefit resulting from their decisions is transferred to a charity caring for ophthalmic patients; for procedural details see Subsection 3.3.

3.2 Payment systems

Recall that each physician decides on the provision of medical services under two different payment systems. In part I of the experiment, physicians decide either under FFS or CAP. In part II, they decide under a mixed payment system. Table 1 provides an overview of the payment systems employed in our experiment. Physicians paid by FFS (CAP) in part I decide under mixed-FFS systems, i.e., Mix-FFS-8 or Mix-FFS-6, (mixed-CAP systems, i.e., Mix-CAP-2 or Mix-CAP-4) in part II. Beyond the within-subject comparison (part I vs. part II), the experimental design also allows us to compare between-subject behavior for FFS and CAP in part I as well as for different mixed payment systems in part II.

[4] Varying patients' characteristics in our lab experiment is motivated by the recent theoretical literature (see, e.g., Allard et al. 2011), which assumes that patient characteristics affect physicians' behavior. The relevance of the patient health benefits is also emphasized by a recent empirical study by Clemens and Gottlieb (2014) who investigate the impact of changes in reimbursement rates on patients' health benefit. They report some differences between younger and more elderly patients the latter likely to be of high severity types. For example, an increase in the reimbursement rate decreases the mortality of younger cardiac patients within 4 years, while older cardiac patients (of age 75 or above) face an increase in mortality within 4 years (see Clemens and Gottlieb, 2014, p.1345).

[5] Note that this is an essential design feature differing from that of Hennig-Schmidt et al. (2011), which does not allow for a systematic analysis of the effects of overprovision and underprovision.

Table 1: Payment systems of the experiment

	Components				
Payment system	μ	L	$1-\mu$	p	R
FFS	–	–	1.00	2	$2q$
CAP	1.00	10	–	–	10
Mix-FFS-8	0.20	18	0.80	2	$3.6+1.6q$
Mix-FFS-6	0.40	16	0.60	2	$6.4+1.2q$
Mix-CAP-2	0.96	10	0.04	10	$9.6+0.4q$
Mix-CAP-4	0.84	10	0.16	5	$8.4+0.8q$

Notes: This table shows the components of the payment systems employed in our experiment. μ is the weight of the lump-sum L and $1-\mu$ is the weight of the fee-for-service component p. R defines the composed remuneration.

The profit functions of FFS and mixed-FFS systems mirror those of the respective CAP and mixed-CAP systems. While varying the components of the payment systems, we keep maximum profit levels and marginal profits for Pareto-efficient choices constant. In FFS, physicians are paid a fee of $p = 2$ per service and $\mu = 0$. Accordingly, profit is $\pi(q) = 2q - q^2/10$. In CAP, physicians receive a lump-sum payment per patient independent of the quantity of medical services $L = 10$. Physicians' profit per patient is thus $\pi(q) = 10 - q^2/10$. The maximum achievable profit is 10 in both FFS and CAP. The profit-maximizing quantity of medical services is 10 and 0 in FFS and CAP, respectively. Also, absolute values of marginal profits are equal in CAP and FFS. The profit parameters are illustrated in Figure 1, and the complete set of parameter values are shown in Table 1 in AppendixA.2.

Mixed payment systems comprise both a lump-sum and a fee-for-service component. In mixed-FFS systems (i.e., Mix-FFS-8 and Mix-FFS-6), a higher weight is attached to the latter. To ensure equality of maximum profits compared to FFS and CAP, we adjust the lump-sum component L. Symmetrically, in mixed-CAP systems (i.e., Mix-CAP-2 and Mix-CAP-4), a larger weight is attached to the lump-sum component. To ensure equal maximum profits, we adjust the fee-for-service component p (see Table 1 for the parameter values).

In sum, mixed payment systems are designed such that incentives inherent in FFS (CAP) to provide too many (few) services are mitigated. Moreover, we choose the profit-maximizing quantities in mixed-CAP systems and mixed-FFS systems to be 2 and 4 as well as 6 and 8, respectively. This ensures that profit-maximizing quantities are 'closer' to the patient-optimal quantities than in non-blended payment systems, but do not coincide with them. We, thus, reduce the trade-offs between profit-maximization and patient health benefit optimization.

We also consider several experimental control conditions. First, keeping the maximum profit constant across FFS, CAP, and mixed payment systems comes at the cost of adjusting the fee-for-service and lump-sum component. Yet retaining the components of the FFS and CAP and using the same profit maximizing quantities as in the mixed payment systems would imply lower maximum profits. In order to control for the behavioral impact associated with this effect, we implement two additional mixed systems, one with more weight on CAP and another with more weight on FFS, using $\mu = 0.20$ or $\mu = 0.80$ but non-adjusted $L = 10$ and $p = 2$. Therefore, $\pi(q) = 2 + 1.6q - q^2/10$ in the former and $\pi(q) = 8 + 0.4q - q^2/10$ in the latter. Second, we control whether subject's behavior in part II (i.e., under mixed payment systems) is affected by preceding decisions regarding non-blended systems in part I. In this control condition, subjects only decide under one mixed payment system (i.e., Mix-FFS-6 or Mix-CAP-4).

3.3 Experimental protocol

The computerized experiment programmed with z-Tree (Fischbacher, 2007) was conducted at elfe, the Essen Laboratory for Experimental Economics at the University of Duisburg-Essen. Overall, 213 students participated in our experimental sessions. Among those were 32 medical students, of whom 16 each participated in payment conditions 2 (FFS/Mix-FFS-6) and 4 (CAP/Mix-CAP-4). Table 2 provides an overview. All subjects were recruited through the online recruiting system ORSEE (Greiner 2004).

The procedure was as follows: upon arrival, subjects were randomly allocated to

Table 2: Experimental conditions

Condition	Part I of the experiment	Part II of the experiment	Subjects
1	FFS	Mix-FFS-8	24
2	FFS	Mix-FFS-6	40
3	CAP	Mix-CAP-2	22
4	CAP	Mix-CAP-4	40
Control conditions			87

cubicles. They were then given ample time to read the instructions for part I and ask clarifying questions, which were answered in private. Subjects were informed that the experiment consisted of two parts but received detailed instructions for part II only after having finished part I of the experiment. To check for subjects' understanding of the decision task, they had to answer a set of control questions. The experiment did not start unless all subjects had answered the control questions correctly. In each of the two parts of the experiment, subjects subsequently decided on the quantity of medical services for each of the nine patients. The order of patients was randomly determined and kept constant for all subjects and all conditions, i.e., $Bx, Cx, Az, By, Bz, Ay, Cz, Ax, Cy$.

Before making their decision for a specific patient, subjects were informed about their remuneration, their cost and profit, as well as about the patient's benefit for each quantity from 0 to 10. All monetary amounts were given in Taler, our experimental currency, the exchange rate being 1 Taler = 0.08 EUR. The procedure was exactly the same in part II of the experiment. At the end of the experiment—when all subjects had made their decisions—we randomly determined one decision in each part of the experiment to be relevant for a subject's actual payoff and the patient benefit. To this was done to rule our income effects. Subjects were paid in private according to the two randomly determined decisions.

To verify that the money corresponding to the sum of patient benefits in a session was actually transferred, we applied a procedure similar to Hennig-Schmidt et al. (2011) and Eckel and Grossman (1996). To this end, one of the participants was randomly chosen to be a monitor. After the experiment, the monitor verified that the order on the amount of the aggregate benefit was written to the financial department of the University of Duisburg-Essen to transfer the money to the Christoffel Blindenmission. The monetary amount supports surgical treatments of cataract patients in a hospital in Masvingo (Zimbabwe) staffed by ophthalmologists from the charity. Notice that we did not inform the participants that the money was assigned to a developing country (see instructions in Appendix A.3). The order was sealed in an envelope, and the monitor and experimenter then walked together to the nearest mailbox and deposited the envelope. The monitor was paid an additional EUR 5.

Sessions lasted for about 70 minutes. Subjects earned, on average, EUR 12.98. The average benefit per patient was EUR 12.26. In total, EUR 2,610 were transferred to the Christoffel Blindenmission. Average costs for a cataract operation amounted to about EUR 30. Thus, our experiment allowed treating 87 patients.

4 Behavioral predictions

In this section, we derive predictions for physicians' behavior in the non-blended and mixed payment systems employed in our experiment. To this end, we use a variant of Ellis and McGuire's (1986) seminal model of a physician deriving utility from her own profit and patients' health benefit.

Let physician i choose the quantity of medical services q in order to maximize her utility

$$U_i(q) = (1 - \alpha_i)\pi(q) + \alpha_i B(q), \tag{3}$$

with $\alpha_i \in [0,1]$. α_i is a measure for physician i's altruism. Hence, for a purely profit-maximizing physician, $\alpha_i = 0$. Based on the specifications of physicians' profits and the patient benefits (i.e., equations 1 and 2) employed in our experiment, we state the following proposition regarding physician i's behavior:

Proposition 1. Physician i's quantity choice q decreases with the weight on the lump-sum component μ. Physician i chooses

$$\begin{aligned} q > q^* & \quad \text{if} \quad (1\text{-}\mu)p > q^*/5 + [\alpha_i/(1-\alpha_i)]\,\theta, \\ q < q^* & \quad \text{if} \quad (1\text{-}\mu)p < q^*/5 - [\alpha_i/(1-\alpha_i)]\,\theta, \\ q = q^* & \quad \text{otherwise.} \end{aligned}$$

Proof. For a formal proof of Proposition 1 see Appendix A.3. □

Proposition 1 indicates that the physician's behavior is influenced by the fee-for-service component $(1-\mu)p$, the patient-optimal quantity of medical services q^*, and the patient's marginal health benefit θ. Also, due to symmetry of the conditions, overprovision and underprovision of medical services decrease in α_i to the same extent. According to the conditions given in Proposition 1, we formulate the following three behavioral hypotheses.

First, we focus on the impact of μ in FFS and CAP compared to mixed payment systems. For the mixed systems, μ increases in the following order: FFS, Mix-FFS-8, Mix-FFS-6, Mix-CAP-4, Mix-CAP-2, and CAP (recall Table 1). According to Proposition 1, quantity gradually decreases in μ independent of the physicians' degree of altruism. Compared to FFS, physicians should reduce overprovision of medical services in the mixed-FFS systems, which should lead to a higher health benefit for patients. As μ is higher in Mix-FFS-6 than in Mix-FFS-8, we expect a higher reduction of overprovision in the former than in the latter. Analogously, in the mixed-CAP systems physicians should reduce underprovision compared to CAP, which again implies a higher patient health benefit. In a similar vein, as μ is higher in Mix-CAP-2 than in Mix-CAP-4, we expect a lower reduction of underprovision in the former than in the latter. We state the following:

Hypothesis 1. Physicians' quantity choices decrease with increasing weight on the lump sum. Overprovision (underprovision) is reduced in Mix-FFS-8 (Mix-CAP-2) and even more so in Mix-FFS-6 (Mix-CAP-4) compared to FFS (CAP).

Second, Proposition 1 indicates that patients' characteristics—the patient-optimal quantity q^* that varies with the severities of illness l, and the patients' marginal health benefit θ that varies with illnesses $k = A, B$ and $k = C$—affect the behavior of at least partially altruistic physicians, i.e., those characterized by $\alpha_i > 0$. Also, from Proposition 1 follows that the level of patient health benefit does not affect behavior. Based on our systematic variation of the patient optimal quantity q^* (severities of illness) and the patient's marginal health benefits (illnesses), we formulate:

Hypothesis 2. Quantity choices of (at least partially altruistic) physicians are affected by patients' characteristics. Physicians' supply of medical services increases with the severity of patients' illness and patients' marginal health benefits.

Finally, Proposition 1 states that overprovision and underprovision of medical services decreases with increasing physician altruism, defined as the weight attached to the patient's health benefit. Our behavioral data allow us to analytically derive the individual physician's degree of altruism based on actual quantity choices. Similar to Godager and Wiesen (2013), who report substantial heterogeneity in physician altruism, we also expect a variation in physician altruism and thus state the following:

Hypothesis 3. Physician altruism is heterogeneous across individuals.

5 Results

5.1 Aggregate provision behavior

For starters, we analyze the provision of medical services at an aggregate level. Table 3 shows descriptive statistics for the payment systems. We find that physicians do respond to the incentives in FFS and CAP. Overall, about 53% less medical services are provided in the latter than in the former. This difference is highly significant ($p = 0.000$, two-sided Mann-Whitney U-Test)[6] and in line with findings reported in earlier empirical and experimental studies (e.g., Gaynor and Gertler 1995; Hennig-Schmidt et al. 2011).

Behavioral data evidence overprovision of medical services in FFS and underprovision in CAP. Overall, $q - q^*$, the deviation of the chosen from the patient-optimal quantity, is 2.11 (s.d. 2.10) services in FFS and -1.67 (s.d. 1.87) services in CAP. These deviations are significant in both payment systems ($p < 0.001$, one-sample Wilcoxon signed-rank test). Nevertheless, physicians do not always choose the profit-maximizing quantity, but take patient benefits into account, at least to some degree. The symmetric design of patient health benefits allows us to test whether incentives to underprovide in CAP are

[6] p-values are from two-sided tests if not indicated otherwise. We neither find any significant differences for FFS nor for CAP in Part I across the two respective sessions ($p > 0.159$, Mann-Whitney U-Test). We, therefore, pool the data and base all tests for FFS and CAP on this data set. Further, notice that non-parametric Fisher-Pitman permutation tests for independent and paired samples corresponding to Mann-Whitney U and Wilcoxon matched-pairs signed rank-tests, respectively, yield comparable p-values. This also applies to parametric t-tests.

Table 3: Aggregate quantities and patient health benefits by payment system

	Part *I* of the experiment				Part *II* of the experiment				
	Payment system	Mean	s.d.	N	Payment system	Mean	s.d.	N	p-values
q	FFS	7.11	2.09	576	Mix-FFS-8	6.92	1.73	216	0.008
					Mix-FFS-6	5.50	1.28	360	0.000
	CAP	3.33	1.95	558	Mix-CAP-2	3.66	1.74	198	0.004
					Mix-CAP-4	4.67	1.35	360	0.000
$B(q)$	FFS	7.51	3.52	576	Mix-FFS-8	7.68	3.38	216	0.011
					Mix-FFS-6	9.44	2.89	360	0.000
	CAP	8.02	3.33	558	Mix-CAP-2	8.27	3.24	198	0.003
					Mix-CAP-4	9.68	2.93	360	0.002

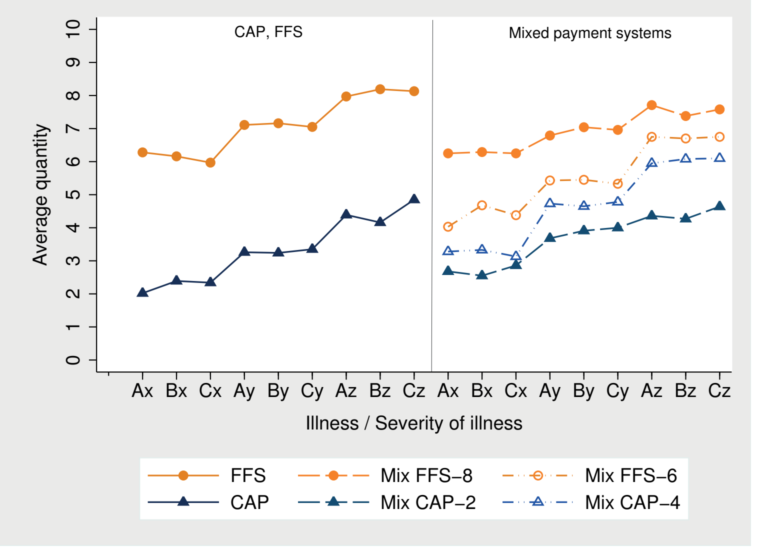
Notes: This table shows descriptive statistics for quantities and patient health benefits for all payment systems. p-values of a Wilcoxon (matched-pairs) signed-rank test are shown for a within-subject comparison of individual service provision across payment systems in both parts of the experiment. N is the number of choices (patients) in a specific treatment.

equally strong as incentives to overprovide in FFS. Even though the extent of non-optimal service provision seems to be slightly higher in CAP compared to FFS, the patient health benefit is not significantly different ($p = 0.216$, Mann Whitney U-Test).

The within-subject comparison reveals that mixed payment systems influence provision behavior as stated in Hypothesis 1. This is illustrated in Figure 1, which shows average provided quantities per patient for non-blended and mixed payment systems. In both Mix-FFS-8 and Mix-FFS-6, physicians deliver significantly less medical services compared to FFS ($p \leq 0.012$, Wilcoxon signed-rank test, see Table 3). In contrast, in Mix-CAP-2 and Mix CAP-4, significantly more medical services are provided than in CAP ($p \leq 0.004$). As a result, we find less overprovision in mixed-FFS systems and less underprovision in mixed-CAP systems as well as an increase in health benefit compared to the non-blended systems ($p \leq 0.012$). In line with Hypothesis 1, overprovision decreases as more weight is attached to the lump-sum component in the mixed-FFS systems ($p = 0.012$, Mann-Whitney U-test). Similarly in the mixed-CAP systems, underprovision decreases with a decreasing lump-sum component, i.e., with a decreasing degree of supply-side cost-sharing ($p = 0.007$, Mann-Whitney U-test). In all mixed systems, we still find a significant deviation from the patient-optimal quantity of medical services ($p \leq 0.004$, Wilcoxon signed rank test).

Next, we investigate the impact of the payment system on patients' health benefit. Health benefit in Mix FFS-8 and Mix FFS-6 is significantly higher than in FFS ($p \leq 0.012$); see again Table 3. The same holds when comparing Mix-CAP-2 and Mix-CAP-4 to CAP ($p < 0.001$). These results imply that a larger weight on the lump-sum component in

Figure 1: Average medical services per illness and severity



Notes: This figure shows average quantities for all patients (sorted by illnesses and severities of illness) by payment systems. The left panel displays average quantities for the non-blended payment systems FFS and CAP. The right panel shows average quantities for the mixed payment systems Mix-FFS-8, Mix-FFS-6, Mix-CAP-2, and Mix-CAP-4.

mixed-FFS systems and a reduction in supply-side cost sharing in mixed-CAP systems increase the patients' health benefit. Patients' health benefit in Mix-FFS-6 (Mix-CAP-4) is significantly higher than in Mix-FFS-8 (Mix-CAP-2) ($p \leq 0.005$, Mann-Whitney U-test). These observations complement Clemens and Gottlieb (2014), who report that patients health is affected by a variation in the payment system.

A regression analysis on the deviation from the patient-optimal quantity $q - q^*$ provides further support for the hypothesis that quantity choices decrease in the weight on the lump-sum component. Panel A of Table 4 shows regression results for mixed-FFS systems, panel B gives results for mixed-CAP systems. FFS and CAP are the respective reference categories. Model (1) comprises dummies for the payment systems Mix FFS-8 and Mix FFS-6, model (4) for the payment systems Mix CAP-2 and Mix CAP-4. In addition, we control for patient health benefit and physician profit, which are important for subjects' choices. Overprovision of medical services is significantly reduced under both mixed-FFS systems compared to FFS. Estimates for coefficients of the mixed-CAP systems in model (4) indicate that underprovision is significantly lower than in CAP. Further, Wald test results reveal that an increase in the lump-sum component from Mix-FFS-8 to Mix-FFS-6 leads to a significant decrease in overprovision of medical services (see Table 4). More-

Table 4: OLS-regressions of overprovision and underprovision of medical services

	A. Mixed-FFS systems			**B. Mixed-CAP systems**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	$q - q^*$	$q - q^*$	$q - q^*$	$q - q^*$	$q - q^*$	$q - q^*$
Mix-FFS-8	-0.617***	-0.677***	-0.701***			
	(0.133)	(0.132)	(0.131)			
Mix-FFS-6	-1.611***	-1.567***	-1.641***			
	(0.075)	(0.077)	(0.093)			
Mix-CAP-2				0.793***	0.811***	0.824***
				(0.191)	(0.185)	(0.181)
Mix-CAP-4				1.674***	1.664***	1.701***
				(0.146)	(0.150)	(0.144)
Patients' health benefit	-0.314***	-0.310***	-0.310***	0.230***	0.230***	0.230***
	(0.019)	(0.019)	(0.019)	(0.027)	(0.027)	(0.027)
Profit	0.556***	0.547***	0.544***	-0.572***	-0.572***	-0.571***
	(0.035)	(0.036)	(0.037)	(0.104)	(0.105)	(0.105)
Med. student		-0.269***	-0.373***		0.071	0.121
		(0.064)	(0.099)		(0.091)	(0.143)
Mix FFS-6 x Med. stud.			0.231**			
			(0.103)			
Mix CAP-4 x Med. stud.						-0.111
						(0.135)
Constant	-0.391	-0.266	-0.216	1.356	1.335	1.321
	(0.282)	(0.301)	(0.315)	(1.010)	(1.022)	(1.032)
Observations	1,152	1,152	1,152	1,116	1,116	1,116
R^2	0.607	0.610	0.610	0.519	0.52	0.52

Notes: The table displays coefficients from ordinary least square-regression models; clustered for individual subjects. The reference category is FFS in panel A and CAP in panel B. Robust standard errors are shown in parentheses under the coefficients. Wald tests show that coefficients for Mix-FFS-8 and Mix-FFS-6 as well as Mix-CAP-2 and Mix-CAP-4 differ significantly ($p = 0.000$). Using alternative estimation techniques, like ordered probit and ordered logit-models, we arrive at very similar results.

***Significant at the 1-percent level.

**Significant at the 5-percent level.

*Significant at the 10-percent level.

over, a decrease of the lump-sum component from Mix-CAP-2 to Mix-CAP-4 leads to a significant decrease of underprovision. In sum, we state the following:

Result 1. *Physicians' quantity choices decrease with increasing weight on the lump-sum component. Overprovision is significantly reduced in mixed-FFS systems, as is underprovision in mixed-CAP systems. In mixed-FFS systems (mixed-CAP systems), a higher share of CAP (FFS) leads to a further reduction in overprovision (underprovision) of medical services.*

Our regressions also control for possible behavioral differences between medical and non-medical students, as some experimental studies indicate that a student's major can have an impact on behavior. For example, Ahlert et al. (2012) find behavioral differences between students of medicine and economics. Hennig-Schmidt and Wiesen (2014) show more other-regrading behavior of medical students compared to non-medical students. In regression models (2), (3), (5) and (6), we control for students' major in medicine (see Table 4), non-medical students being the reference category. We find that our results are robust with regard to subject pool differences. For medical students, the reduction in overprovision of medical services is more pronounced in mixed-FFS systems. For mixed-CAP systems, medical students' behavior is not significantly different from that of other majors.

In our experimental control conditions, we investigate, first, whether the adjustment of the lump sum L and fee p in mixed payment systems and, second, whether 'subjects' experience' based on the non-blended systems in part I affect our results. When comparing behavior in mixed payment systems with adjusted fee-for-service and the lump-sum components (i.e., Mix-FFS-8 and Mix-CAP-2) with corresponding mixed systems with reduced profit levels, we do not find significant differences ($p \geq 0.224$, Mann-Whitney U-Test). When controlling for experience we also do not find significant differences ($p \geq 0.205$, Mann-Whitney U-Test).

5.2 Influence of patients' characteristics

We now investigate how patient characteristics affect physicians' behavior according to Hypothesis 2. In particular, we analyze the effects of the patient-optimal quantities (q^*) and the patients' marginal health benefits (θ). The former vary with the severities of illness: $q^* = 3, 5, 7$ for x, y, and z, respectively. The latter vary with illnesses: for A and B, $\theta = 1$ and for C, $\theta = 2$.

Figure 1 illustrates that the severities of illnesses systematically affect physicians' quantity choices. In line with Hypothesis 2, average quantities increase with the severity of illness in all payment systems (descriptive data and p-values are shown in Panel B of Table

2 in Appendix A.2). We observe the following pattern: lowest quantities are provided for the low severity x, medium quantities for the intermediate severity y, and highest quantities for severity z. Pairwise comparisons indicate significant differences across severities of illnesses for all payment systems ($p \leq 0.012$, t-test for paired samples).

The maximum health benefit (varied with illnesses k) and the marginal health benefit do not greatly affect physicians' quantity choices (see Panel A of Table 2 in Appendix A.2). Comparing how patients with illnesses A and B are treated reveals significant differences for Mix-FFS-6 only. That is, in most cases the maximum patient health benefit does not affect behavior. When comparing quantity choices for illnesses B and C, i.e., when varying both the maximum and the marginal health benefit, we find only significant differences for CAP and Mix-FFS-6. In both conditions, physicians choose significantly more medical services for illness C compared to illness B. Pairwise comparisons for the remaining payment systems indicate no significant differences ($p \geq 0.222$). In sum, we state:

Result 2. *Quantities of medical services significantly increase with the severity of illness in all payment systems. Marginal health benefits do not affect physicians' quantity choices much.*

5.3 Heterogeneity in physician altruism

In the following, we investigate the degrees of physician altruism—the weights attached to patients' health benefit. We infer the individual degree of altruism from a physician's actual quantity choices for the 18 patients, nine in part I and nine in part II of the experiment.

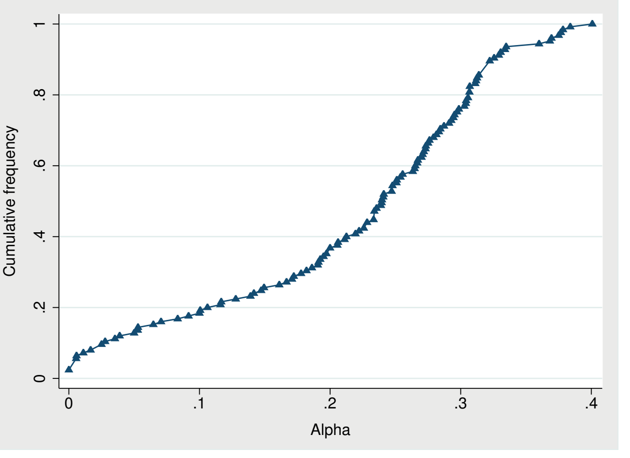
Recall that each of the physician's quantity choices determines a profit/patient health benefit pair. Using equation (3), this allows us to calculate a physician's utilities for patient j for different levels of $\alpha_i \in [0,1]$. Given physician i's quantity choice for patient j, we calculate α_{ij} which maximizes a physician i's utility.[7] Averaging over 18 quantity choices in non-blended FFS, CAP, and mixed payment systems then gives us the individual physician i's α_i.

We find that physicians attach a positive weight to the patient's health benefit. On the aggregate, $\overline{\alpha}_i = 0.22$ (s.d. 0.18).[8] Figure 2 plots the distribution of physician altruism in our experimental sample, indicating considerable heterogeneity, as demonstrated by

[7] In the mixed payment system Mix-CAP-4, for example, a physician's quantity choice of $q = 4$ for patient Cz implies $\alpha = 0$. A choice of $q = 5$ implies that a physician's utility is maximized at $\alpha = 0.1$. For a physician choosing the patient optimal quantity ($q = 7$), utility is maximized for all $\alpha \geq 0.2$. Here, we then select the minimum α still maximizing a physician's utility at $q = 7$, being $\alpha = 0.2$ in our example.

[8] We include all subjects into our analysis who made at least 12 Pareto-efficient choices. For that reason,

Figure 2: Distribution of physician altruism



Notes: This figure illustrates the cumulative frequency of individuals' average physician altruism in our experimental data. α is calculated for each physician and averaged over all 18 decisions of a physician. Only those of the 126 subjects in the main experimental conditions were included who made at least 12 Pareto-efficient quantity choices ($N = 125$).

Godager and Wiesen (2013). Table 5 indicates the same result. In particular, around 14 % of physicians attach a low weight of $\alpha_i \leq 0.05$ to the patients' health benefit. Such a low α implies choosing the profit-maximizing quantity for most patients and never providing the patient-optimal quantity of medical services. These subjects can be characterized as rather profit-maximizing. For about 24 % of individuals, $0.05 < \alpha_i \leq 0.20$. For most subjects, namely 40 %, $0.2 < \alpha_i \leq 0.3$, meaning that they put a considerable weight on the patients' health benefit. About 22 percent of subjects attach a weight larger than 0.3 to the patient's health benefit, which implies the choice of the patient-optimal quantity in all mixed payment systems.

Table 5 shows a substantial share of non-medical students who are profit-maximizers ($\alpha_i \leq 0.05$) compared to medical students which implies some variation in this category across subject pools. Yet the difference is not significant ($\overline{\alpha}^{Med} = 0.24$ (s.d. 0.18) and $\overline{\alpha}^{Non-Med} = 0.21$ (s.d. 0.18), $p = 0.2450$, Mann-Whitney U-test. In sum, we state the following result.

Result 3. *There is substantial heterogeneity in the individual degree of physician altruism.*

---

one of the 126 subjects who participated in one of the four conditions FFS/Mix-FFS and CAP/Mix-CAP had to be excluded.

Table 5: Heterogeneity in physician altruism

Category	All	Med. students	Non-med. students
$\alpha_i \leq 0.05$	0.14	0.03	0.18
$0.05 < \alpha_i \leq 0.20$	0.24	0.18	0.26
$0.20 < \alpha_i \leq 0.30$	0.40	0.58	0.33
$\alpha_i > 0.30$	0.22	0.21	0.23
N	125	32	93

Notes: This table shows relative frequencies for different categories of α_i for all subjects and for medical and non-medical students separately.

6 Concluding remarks

We use laboratory experiments to investigate the causal effects of introducing mixed payment systems, which comprise a fee-for-service and a lump-sum component and vary in the weighting of components, on the provision of medical services. Prior to the mixed systems, subjects' decisions are paid through non-blended fee-for-service or capitation systems. Patient characteristics are varied systematically and kept constant for all payment systems. Our results are consistent with behavioral predictions derived from a variant of Ellis and McGuire's (1986) seminal model of physician behavior.

We find that medical service provision decreases with increasing weight on the lump-sum component. Overprovision and underprovision inherent in fee-for-service and capitation systems, respectively, are significantly reduced by introducing mixed payment systems which leads to higher patient health benefits. Decisions are significantly influenced by the patient-optimal quantities: the more medical services a patient needs to achieve his maximum health benefit (i.e., the higher the severity of illness), the higher the quantity provided. Marginal health benefits have only a minor impact on behavior.

The observed effect of physicians' quantity choices are largely in line with (partly) altruistic preferences. While the aggregate results indicate a positive degree of physician altruism, individual degrees vary substantially. These results suggest, according to Ellis and McGuire (1986), that while some individuals behave similar to a 'perfect agent', it is necessary to account for imperfect agency in models of physician behavior. In particular, as non-blended FFS leads to significant deviations from patient-optimal behavior, mixed payment systems with an element of supply-side cost sharing have the potential to move medical service provision towards the patient optimum.

The heterogeneity in physician altruism would make it natural to offer physicians a menu of payment systems, as opposed to a 'one-size-fits-all' system. Several theoretical papers argue, for example, that the more altruistic physicians should be paid on a capitated basis (e.g., Eggleston, 2005, Jack, 2005) regardless of the patients severity of illness.

Barham and Milliken (forthcoming) show, however, that high-severity patients should be treated by physicians paid on a fee-for-service basis, whereas the least altruistic physicians treating healthy patients should be paid by capitation. This is consistent with our behavioral results for patients with different severities of illnesses. We demonstrate that mixed payment systems—if designed such that tradeoffs between profit-maximization and rendering the optimal benefit to the patient are reduced—might help to incentivize less altruistic physicians to choose patient-optimal care for both low and high-severity patients.

In terms of future research, physicians' self-selection into payment systems requires attention. To offer physicians the 'right' menu of payment schemes, it is, for example, important to better understand physicians' underlying motivations driving their decision in this respect. Also, the impact transitions in patients' health status—for example, from low to high severity, in light of the increasing number of elderly and frail patients—have on physicians' behavior in different payment systems seems to be a relevant research topic.

References

Ahlert, M., S. Felder, and B. Vogt (2012): "Which patients do I treat? An experimental study with economists and physicians," *Health Economics Review*, 2, doi:10.1186/2191–1991–2–1.

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

Arrow, K. J. (1963): "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53, 941–969.

Baicker, K. and D. Goldman (2011): "Patient Cost-Sharing and Healthcare Spending Growth," *Journal of Economic Perspectives*, 25, 47–68.

Barham, V. and O. Milliken (forthcoming): "Payment mechanisms and the composition of physician practices: balancing cost-containment, access, and quality of care," *Health Economics*.

Buckley, N. J., K. Cuff, J. Hurley, L. McLeod, S. Mestelman, and D. Cameron (2012): "An experimental investigation of mixed systems of public and private health care finance," *Journal of Economic Behavior Organization*, 84, 713 – 729.

Chalkley, M. and J. M. Malcomson (1998): "Contracting for Health Services when Patient Demand Does not Reflect Quality," *Journal of Health Economics*, 17, 1–19.

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

Clemens, J. and J. D. Gottlieb (2014): "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104, 1320–1349.

Dumont, E., B. Fortin, N. Jacquement, and B. Shearer (2008): "Physicians' Multitasking and Incentives: Empirical Evidence from a Natural Experiment," *Journal of Health Economics*, 27, 1436–1450.

Eckel, C. and P. Grossman (1996): "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 16, 181–191.

Eggleston, K. (2005): "Multitasking and mixed payment systems for provider payment," *Journal of Health Economics*, 24, 211–223.

Ellis, R. P. and T. G. McGuire (1986): "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, 5, 129–151.

——— (1990): "Optimal Payment Systems for Health Services," *Journal of Health Economics*, 9, 375–396.

——— (1993): "Supply-Side and Demand-Side Cost Sharing in Health Care," *Journal of Economic Perspectives*, 7, 135–151.

Falk, A. and J. J. Heckman (2009): "Lab Experiments Are a Major Source of Knowledge in Social Sciences," *Science*, 326, 535–538.

Fischbacher, U. (2007): "Z-tree: Zurich Toolboox for Readymade Economic Experiments – Experimenter's Manual," *Experimental Economics*, 10, 171–178.

Frank, R. (2007): "Behavioral Economics and Health Economics," in *Behavioral Economics and Its Applications*, ed. by P. Diamond and H. Vartianinen, Princeton University Press, 195–234.

Fuchs, V. (2000): "The Future of Health Economics," *Journal of Health Economics*, 19, 141–157.

Gaynor, M. and P. Gertler (1995): "Moral Hazard and Risk Spreading in Partnerships," *Rand Journal of Economics*, 26, 591–613.

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

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments," in *Forschung und wissenschaftliches Rechnen : Beiträge zum Heinz-Billing-Preis 2003*, ed.

by K. Kremer and V. Macho, Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen, 79–93.

Hanson, K. and W. Jack (2010): "Incentives Could Induce Ethiopian Doctors and Nurses to Work in Rural Settings," *Health Affairs*, 29, 1452–1460.

Hennig-Schmidt, H., R. Selten, and D. Wiesen (2011): "How Payment Systems Affect Physicians' Provision Behavior – An Experimental Investigation," *Journal of Health Economics*, 30, 637–646.

Hennig-Schmidt, H. and D. Wiesen (2014): "Other-regarding behavior and motivation in health care provision: An experiment with medical and non-medical students," *Social Science & Medicine*, 108, 156 – 165.

Jack, W. (2005): "Purchasing Health Care Services from Providers with Unknown Altruism," *Journal of Health Economics*, 24, 73–93.

Kantarevic, J., B. Kralj, and D. Weinkauf (2011): "Enhanced fee-for-service model and physician productivity: Evidence from Family Health Groups in Ontario," *Journal of Health Economics*, 30, 99–111.

Kessler, D. and M. McClellan (1996): "Do Doctors Practice Defensive Medicine?" *Quarterly Journal of Economics*, 111, 353–390.

Kesternich, I., H. Schumacher, and J. Winter (2014): "Professional norms and physician behavior: homo oeconomicus or homo hippocraticus?" Discussion Paper Series of SFB/TR 15 Governance and the Efficiency of Economic Systems, No. 456.

Kralj, B. and J. Kantarevic (2013): "Quality and quantity in primary care mixed-payment models: evidence from family health organizations in Ontario," *Canadian Journal of Economics*, 46, 208–238.

Krasnik, A., P. Groenewegen, P. Pedersen, P. von Scholten, G. Mooney, A. Gottschau, H. Flierman, and M. Damsgaard (1990): "Changing Remuneration Systems: Effects on Activity in General Practice," *British Medical Journal*, 300, 1698–1701.

Liu, T. and C. Ma (2013): "Health Insurance, Treatment Plan, and Delegation to Altruistic Physician," *Journal of Economic Behavior & Organization*, 85, 79 – 96.

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.

Rizzo, J. A. and R. J. Zeckhauser (2007): "Pushing Incomes to Reference Points:

Why do male Doctors Earn More?" *Journal of Economic Behavior & Organization*, 63, 514–536.

Schram, A. and J. Sonnemans (2011): "How Individuals Choose Health Insurance: An Experimental Analysis," *European Economic Review*, 55, 799–819.

Scott, A. and P. Sivey (2013): "Measuring the Monetary Motivation of Physicians," *Working Paper, University of Melbourne.*

Siciliani, L. (2009): "Paying for Performance and Motivation Crowding Out," *Economics Letters*, 103, 68 – 71.

Yip, W. C. (1998): "Physician response to Medicare fee reductions: changes in the volume of coronary artery bypass graft (CABG) surgeries in the Medicare and private sectors," *Journal of Health Economics*, 17, 675 – 699.

A Appendix

A.1 Instructions

Welcome to the Experiment!

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 the decisions you make. 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. 10 Taler equals 8 Euro. The experiment will take about 90 minutes and consists of two parts. You will receive detailed instructions before each part. Note that none of your decisions in either part have any influence on the other part of the experiment.

Part I of the experiment

Please read the 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. Part I of the experiment consists of 9 rounds of decision situations.

Decision situations

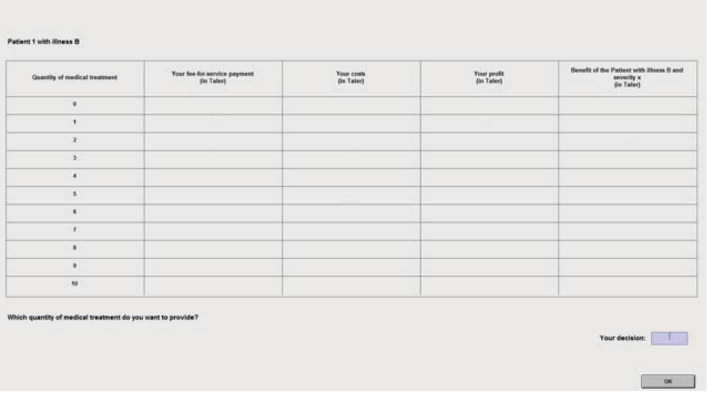
In each round, you are in the role of a physician and decide on medical treatment for a patient. That is, you determine the quantity of medical services you wish to provide to the patient for a given illness and a given severity of this illness. Each patient is characterized by one of three illnesses (A, B, C), each of which can occur in three different degrees of severity (x, y, z). In each consecutive decision round you will face one patient who is characterized by one of the 9 possible combinations of illnesses and degrees of severity (in random order). Your decision is to provide each of these 9 patients with a quantity of $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, or 10 medical services.

Payment

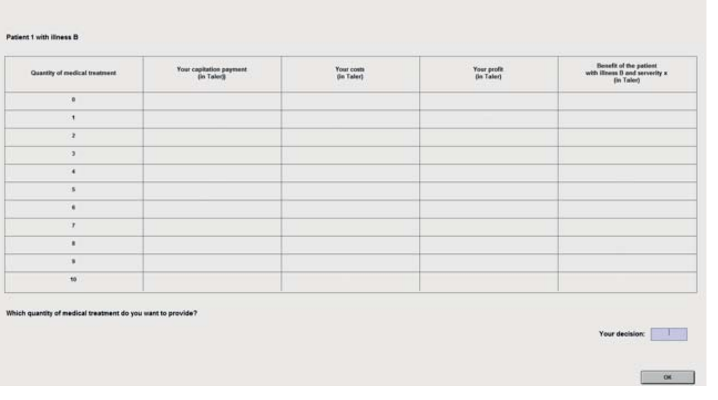
In each round you receive a fee-for-service (capitation) remuneration for treating the patient. Your remuneration increases with the amount of medical treatment (irrespective of the amount of medical treatment) you provide. You also incur costs for treating the patient, which likewise depend on the quantity of services you provide. Your profit for each decision is calculated by subtracting these costs from the fee-for-service (capitation) remuneration. Each quantity of medical service yields a particular benefit for the patient—contingent on his illness and severity. 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) for the respective patient, the illness, your amount of fee-for-service (capitation) remuneration—for each possible amount of medical treatment—your costs, profit, as well as the benefit for the patient with the corresponding illness and severity.

Screen in FFS



Screen in CAP



Payoff

At the end of the experiment, one of the 9 rounds in part *I* will be chosen at random. Your profit in this round will be paid to you in cash.

For this part of the experiment, no patients are physically present in the laboratory. Yet the patient benefit does accrue to a real patient: The amount resulting from your decision will be transferred to the Christoffel Blindenmission Deutschland e.V., 64625 Bensheim, an organization which funds the treatment of patients with eye cataracts.

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 in the randomly chosen situation. This form prompts the payment of the designated amount to the Christoffel Blindenmission Deutschland e.V. by the finance department of the University of Duisburg-Essen. The form is then sealed in a stamped envelope and deposited 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 that 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. Part I of the experiment will begin once all participants have answered the comprehension questions correctly.

Part II of the experiment

Please read the 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. Part II of the experiment also consists of 9 rounds of decision situations.

Decision Situations

As in Part I of the experiment, you take on the role of a physician in each round and decide on medical treatment for a patient. That is, you determine the quantity of medical services you wish to provide to the patient for a given illness and a given severity of this illness.

Each patient is characterized by one of three illnesses (A, B, C), each of which can occur in three different degrees of severity (x, y, z). In each consecutive decision round you will face one patient who is characterized by one of the 9 possible combinations of illnesses and degrees of severity (in random order). Your decision is to provide each of these 9 patients with a quantity of $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, or 10 medical services.

Payment

In each round you are remunerated for treating the patient. In each round you receive a fee-for-service (capitation) remuneration for treating the patient. Your remuneration increases with the amount of medical treatment (irrespective of the amount of medical treatment) you provide. In addition to this, in each round you receive a capitation remuneration that is irrespective of the amount of medical treatment (a fee-for-service remuneration which increases with the amount of medical treatment). You also incur costs for treating the patient, which likewise depend on the quantity of services you provide. Your profit for each decision is calculated by subtracting these costs from the sum of your fee-for-service (capitation) and capitation (fee-for-service) remuneration. As in Part I, each quantity of medical service yields a particular benefit for the patient—contingent on his illness and severity. 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) for the respective patient the illness, your amount of fee-for-service (capitation) remuneration—for each possible amount of medical treatment—the amount of your capitation (fee-for-service) remuneration, your costs, profit, as well as the benefit for the patient with the corresponding illness and severity.

Payoff

At the end of the experiment one of the 9 rounds of Part II will be chosen at random. Your profit in this round will be paid to you in cash, in addition to your payment from the round chosen for Part I of the experiment. After the experiment is over, please remain seated until the experimenter asks you to step forward. You will receive your payment at the front of the laboratory before exiting the room.

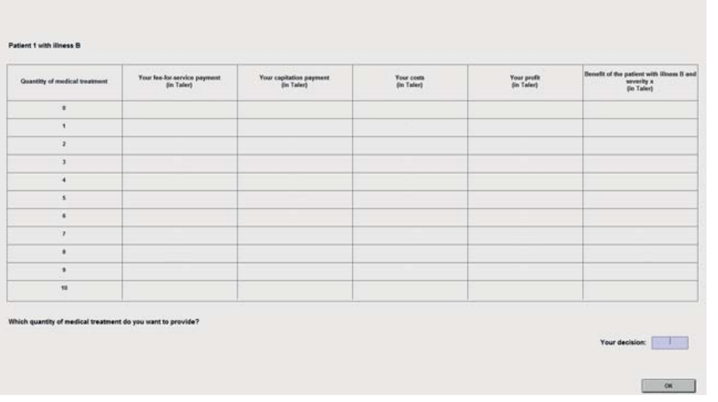
As in Part I, no patients are physically present in the laboratory for Part II of the experiment. Yet the patient benefit does accrue to a real patient: The amount resulting from your decision will be transferred to the Christoffel Blindenmission Deutschland e.V., 64625 Bensheim, an organization which funds the treatment of patients with eye cataracts.

The process for the transfer of money to the Christoffel Blindenmission Deutschland e.V., as described for part I of the experiment, will be carried out by the experimenter and one participant.

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. Part II of the experiment will begin

Screen in Mixed FFS systems



Patient 1 with illness B

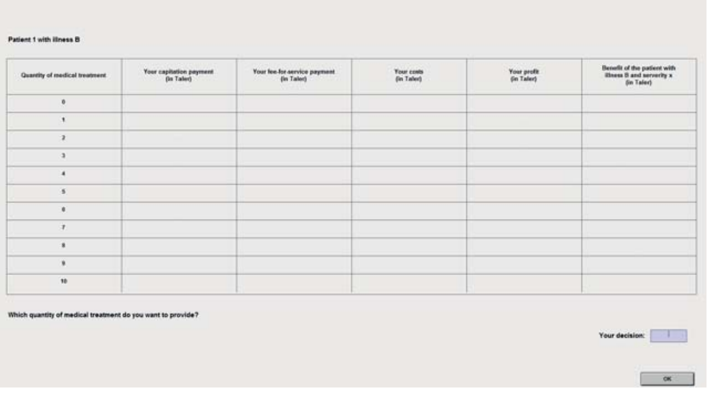
Quantity of medical treatment	Your fee-for-service payment (in Taler)	Your capitation payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Benefit of the patient with illness B and severity x (in Taler)
0					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Which quantity of medical treatment do you want to provide?

Your decision:

OK

Screen in Mixed CAP systems



Patient 1 with illness B

Quantity of medical treatment	Your capitation payment (in Taler)	Your fee-for-service payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Benefit of the patient with illness B and serverity x (in Taler)
0					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Which quantity of medical treatment do you want to provide?

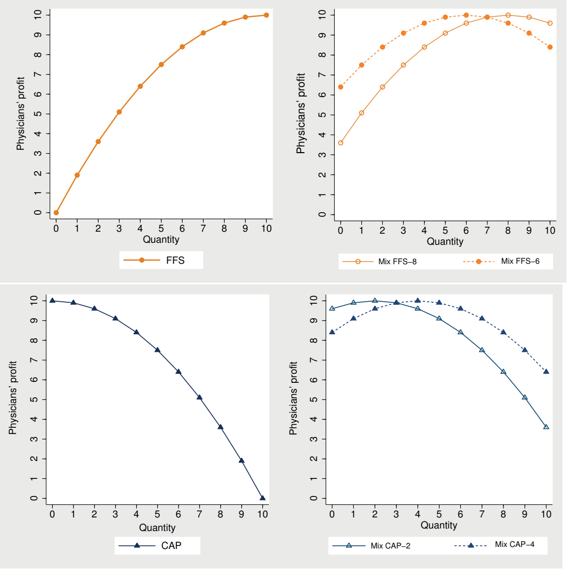
Your decision:

OK

once all participants have answered the comprehension questions correctly. Finally, we kindly ask you to not talk to anyone about the content of this session in order to prevent influencing other participants after you. Thank you for your cooperation!

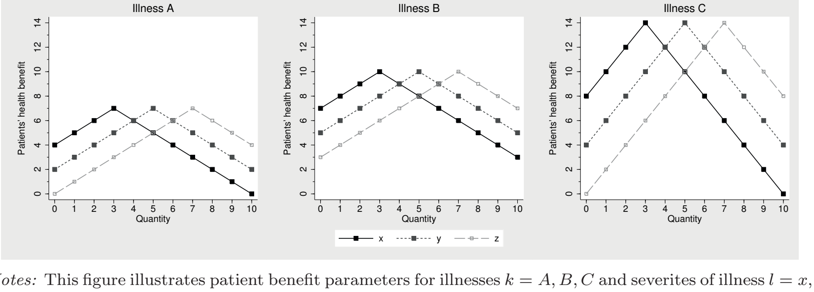
A.2 Additional figures and tables

Figure 1: Profit parameters in CAP, FFS, and mixed payment systems



Notes: This figure illustrates profit parameters in different payment conditions of the experiment. Profits in FFS and CAP (part *I* of the experiment) are shown in the left panel. The right panel illustrates profits in part *II* of the experiment in mixed-FFS systems Mix-FFS-8 and Mix-FFS-6 or mixed-CAP systems Mix-CAP-2 and Mix-CAP-4.

Figure 2: Patient health benefits



Notes: This figure illustrates patient benefit parameters for illnesses $k = A, B, C$ and severites of illness $l = x, y, z$, which are kept constant for all payment conditions.

Table 1: Experimental parameters

		Quantity (q)										
		0	1	2	3	4	5	6	7	8	9	10
Patient	B_{Ax}	4	5	6	7	6	5	4	3	2	1	0
benefit	B_{Ay}	2	3	4	5	6	7	6	5	4	3	2
	B_{Az}	0	1	2	3	4	5	6	7	6	5	4
	B_{Bx}	7	8	9	10	9	8	7	6	5	4	3
	B_{By}	5	6	7	8	9	10	9	8	7	6	5
	B_{Bz}	3	4	5	6	7	8	9	10	9	8	7
	B_{Cx}	8	10	12	14	12	10	8	6	4	2	0
	B_{Cy}	4	6	8	10	12	14	12	10	8	6	4
	B_{Cz}	0	2	4	6	8	10	12	14	12	10	8
Costs	c	0.0	0.1	0.4	0.9	1.6	2.5	3.6	4.9	6.4	8.1	10.0
FFS	p	0.0	2.0	4.0	6.0	8.0	10.0	12.0	14.0	16.0	18.0	20.0
	π	0.0	1.9	3.6	5.1	6.4	7.5	8.4	9.1	9.6	9.9	10.0
CAP	L	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
	π	10.0	9.9	9.6	9.1	8.4	7.5	6.4	5.1	3.6	1.9	0.0
Mix-FFS-8	μL	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6	3.6
	$(1-\mu)p$	0.0	1.6	3.2	4.8	6.4	8.0	9.6	11.2	12.8	14.4	16.0
	π	3.6	5.1	6.4	7.5	8.4	9.1	9.6	9.9	10.0	9.9	9.6
Mix-FFS-6	μL	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4	6.4
	$(1-\mu)p$	0.0	1.2	2.4	3.6	4.8	6.0	7.2	8.4	9.6	10.8	12.0
	π	6.4	7.5	8.4	9.1	9.6	9.9	10.0	9.9	9.6	9.1	8.4
Mix-CAP-2	μL	9.6	9.6	9.6	9.6	9.6	9.6	9.6	9.6	9.6	9.6	9.6
	$(1-\mu)p$	0.0	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2	3.6	4.0
	π	9.6	9.9	10.0	9.9	9.6	9.1	8.4	7.5	6.4	5.1	3.6
Mix-CAP-4	μL	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4
	$(1-\mu)p$	0.0	0.8	1.6	2.4	3.2	4.0	4.8	5.6	6.4	7.2	8.0
	π	8.4	9.1	9.6	9.9	10.0	9.9	9.6	9.1	8.4	7.5	6.4

Notes: This table shows the parameters used in our experiment for all payment conditions. Notice that $(1-\mu)p$ is the weighted FFS-component, μL is the weighted lump-sum component in mixed payment systems, and π is the physician's profit.

Table 2: Influence of patients' characteristics

	A. Illnesses						**B. Severity of illness**					
Payment	Quantity of medical services			Comparisons, p-values			Quantity of medical services			Comparisons, p-values		
system	$A,\ \theta = 1$	$B,\ \theta = 1$	$C,\ \theta = 2$	A vs. B	A vs. C	B vs. C	$x,\ q^* = 3$	$y,\ q^* = 5$	$z,\ q^* = 7$	x vs. y	x vs. z	y vs. z
FFS	7.12	7.17	7.05	0.691	0.437	0.264	6.14	7.10	8.09	0.000	0.000	0.000
	(2.10)	(2.13)	(2.05)				(2.48)	(1.87)	(1.28)			
CAP	3.22	3.26	3.52	0.689	0.022	0.006	2.25	3.28	4.47	0.000	0.000	0.000
	(1.99)	(1.79)	(2.05)				(1.26)	(1.63)	(2.17)			
Mix-FFS-8	6.92	6.90	6.93	0.839	0.867	0.737	6.26	6.93	7.56	0.000	0.001	0.004
	(1.73)	(1.78)	(1.71)				(2.17)	(1.62)	(0.93)			
Mix-FFS-6	5.40	5.61	5.48	0.003	0.096	0.027	4.36	5.40	6.73	0.000	0.000	0.000
	(1.34)	(1.23)	(1.26)				(1.25)	(0.54)	(0.44)			
Mix-CAP-2	3.58	3.58	3.83	0.998	0.326	0.222	2.70	3.86	4.42	0.000	0.000	0.012
	(1.55)	(1.72)	(1.93)				(1.05)	(1.52)	(2.04)			
Mix-CAP-4	4.65	4.68	4.67	0.521	0.725	0.754	3.24	4.72	6.04	0.000	0.000	0.000
	(1.32)	(1.34)	(1.38)				(0.47)	(0.47)	(1.03)			

Notes: This table shows mean quantities of medical services by illnesses (A, B, C) and severities of illness (x, y, z) for the six different payment systems. Standard deviations are shown in parentheses. p-values of two-sided t-tests for paired samples are shown.

A.3 Proof of Proposition 1

The physician's objective function $U(q) = (1-\alpha_i)\pi(q) + \alpha_i B(q)$ is concave as defined in equation (2). Payment $R(q) = \mu L + (1-\mu)pq$, with $\mu > 0$, is linear and $-c(q)$ is concave as $c(q)$ is convex, thus $\pi(q)$ is a concave function. As $B(q)$ is also concave function and $\alpha \geq 0$, $U(q)$ is concave.

Note that as $B(q)$ is not differentiable at $q = q^*$, with $q^* \in (0, 10)$. For $q < q^*$, the first-order condition $U'(q) = (1-\alpha)\left[(1-\mu)p - \frac{q}{5}\right] + \alpha\theta$. For $q > q^*$, the first-order condition $U'(q) = (1-\alpha)\left[(1-\mu)p - \frac{q}{5}\right] - \alpha\theta$.

For $q > q^*$, consider $\lim_{q \to q^*} U'(q) = (1-\alpha)\left[(1-\mu)p - \frac{q^*}{5}\right] - \alpha\theta$. If (1-$\mu$)$p < q^*/5 - [\alpha/(1-\alpha)]\,\theta$, $\lim_{q \to q^*} U'(q)$ is positive. Also because $U(q)$ is concave, $U'(q) > 0 \;\forall\; q < q^*$. Therefore any q such that $q \leq q^*$ cannot be optimal, i.e., physician i chooses $q > q^*$.

Analogously for $q < q^*$, consider $\lim_{q \to q^*} U'(q) = (1-\alpha)\left[(1-\mu)p - \frac{q^*}{5}\right] + \alpha\theta$. If (1-$\mu$)$p < q^*/5 + [\alpha/(1-\alpha)]\,\theta$, $\lim_{q \to q^*} U'(q)$ is negative. Also because $U(q)$ is concave, $U'(q) < 0 \;\forall\; q > q^*$. Therefore any q such that $q \geq q^*$ cannot be optimal, i.e., physician i chooses $q < q^*$.