How Changes in Payment Schemes Influence Provision Behavior

Jian Wang
Department of Health Management and Health Economics, University of Oslo

Tor Iversen
Department of Health Management and Health Economics, University of Oslo

Heike Hennig-Schmidt
Department of Economics, University of Bonn and Department of Health Management and Health Economics, University of Oslo

Geir Godager
Department of Health Management and Health Economics, University of Oslo

UNIVERSITY OF OSLO
HEALTH ECONOMICS RESEARCH NETWORK

Working paper 2017: 2

Health Economics Research Network at the University of Oslo
ISSN 1501-9071 (print version.), ISSN 1890-1735 (online), ISBN 978-82-7756-261-2

How Changes in Payment Schemes Influence Provision Behavior.

Evidence from a lab experiment with physicians and medical students[☆]

Jian Wang [1,4], Tor Iversen [1], Heike Hennig-Schmidt[1,3], Geir Godager*[1,2]

Abstract

When implementing a payment reform, policy makers face the challenge of assessing the effects on health care providers' behavior. Empirical evidence most often relies on field studies, register- or survey data characterized by the absence of a control group. In this paper, we conduct a controlled laboratory experiment to assess the effect of a change in the payment system using the parameters of Hennig-Schmidt, Selten, and Wiesen (2011). We focus on the two payment systems fee-for-service (FFS) and capitation (CAP). Participants are either practicing medical doctors or medical students. They are confronted with two different payment mechanisms, transitioning either from FFS to CAP or vice versa. We also analyze whether the effect of financial incentives is dependent on cultural context by comparing German and Chinese medical students' provision behavior. In line with previous evidence, both doctors and medical students provide fewer medical services under CAP than under FFS. Patient benefit deviates significantly from the patient optimum under both payment systems, even though subjects do not maximize their profits. Whether CAP or FFS is beneficial for the patient depends on the patient type. We find that doctors provide less patient benefit and less frequently choose benefit-maximizing treatments. We find that the sequence of payment schemes affects physician provision behavior. Under CAP, more benefit is provided when CAP follows FFS as compared to the opposite order. Under FFS, we observe no such effect. The interpretation is that provider behavior under a payment schemes can depend not only on the current payment scheme, but also on the payment scheme that was implemented in the past. Comparing medical students from Germany and China, we do not find any difference in behavior under the same payment scheme.

Keywords:

JEL-Classification: C91, I11, H40, J33,

[☆]Financial support by the Research Council of Norway (IRECOHEX, Project-No. 231776) and by National Nature Science Foundation of China (NSFC, Grant-No. 71373146) is gratefully acknowledged.

*Corresponding author. Email: geir.godager@medisin.uio.no

[1]Institute of Health and Society, Department of Health Management and Health Economics, University of Oslo, Norway.

[2]Health Services Research Unit, Akershus University Hospital, Norway

[3]BonnEconLab, University of Bonn, Germany

[4]Center for Health Economic Experiments and Public Policy, School of Public Health, Shandong University, China

1 Introduction

The question of how physicians should be paid in order to promote higher quality health care services while controlling costs has been central in health economics research for decades. Understanding how physicians respond to economic incentives is fundamental when aiming to achieve these goals. The existing theoretical literature and the empirical literature based on field data and from controlled laboratory experiments provide evidence that the design of a payment system for health care providers affects their decisions (see for example Ellis and McGuire, 1986, 1990; Scott, 2000; Gosden et al., 2001; Iversen and Lurås, 2000; Iversen, 2004; Yip et al., 2010; Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2016, 2017a, 2017b). When analyzing the most common forms of physician payment—fee-for-service (FFS) and per-capita payment (CAP) (see, e.g., McGuire, 2000)—a reoccurring result is that the former promotes activity, and the resulting service volume can be higher than optimal. Likewise, the latter prospective payment system encourages the provision of few services, and the resulting service volume tends to be smaller than optimal (Newhouse, 1996).

FFS-based payment systems have traditionally been the prevailing payment method for health care providers in many countries around the world. However, rapidly increasing health care expenditures, which have been attributed to the significant importance of physicians' payment—particularly in the fee-for-service payment schemes—have become a significant challenge (see the discussion in Hennig-Schmidt et al., 2011; Yip and Hsiao, 2008; Eggleston, 2012). In recent years, policy makers in many countries (e.g. USA, China, Germany and Norway) have implemented health care reforms using prospective payment methods including capitation in order to reduce health-related expenditures. When implementing a payment reform, policy makers, however, face the challenge of accounting for health care providers' behavior. Knowledge of responses to changes in economic incentives is necessary when calibrating payment systems to target desired results. The empirical evidence in the literature on how payment schemes affect physician behavior most often relies on field studies, register- or survey data. These data are characterized by an absence of control, which is necessary in order to provide reliable causal inferences about the effects of incentives. Uncontrolled variations in the field can include e.g. unobserved characteristics of the patient population or self-selection of providers (Gaynor and Gertler, 1995; Sørensen and Grytten, 2003; Grytten et al., 2009; Devlina and Sarma, 2008).

Laboratory experiments—being a complementary approach to surveys, field studies, randomized control trials (RTC), and payment-reform experiments in the field—have the unique feature of allowing the researcher to investigate the causal effects of a change in the payment system (a reform) on behavior, since laboratory experiments provide ceteris-paribus variations: The payment system is manipulated while all other variables are kept constant (Falk and Fehr, 2003; Falk and Heckman, 2009).. Controlled lab experiments also have great potential as a 'test bed' for field experiments, large-scale studies, and institutional changes of the health care market before these changes are implemented. They require much less time and financial resources to be implemented and analyzed (see Hennig-Schmidt et al., 2011; Cox et al., 2016a). Finally, laboratory experimentation provides a scalable approach, as it allows for the flexible adaptation of the experimental setting.

We conduct a controlled laboratory experiment, designed to identify the causal effect of payment schemes on provider behavior. We focus on the two non-blended payment systems FFS and CAP. Participants are Chinese medical doctors as well as Chinese medical students expected to become physicians in the future. In our within-subject-design, we study a "reform" by letting subjects be confronted sequentially with two different payment schemes, transitioning either from FFS to CAP or vice versa. We apply experimental parameters identical to those in Hennig-Schmidt et al. (2011) and use a medically framed setting in which subjects in the role of physicians make decisions on the provision of medical services. A subject's quantity choice determines his or her own profit and a patient's health benefit. Decisions are incentivized by monetary rewards determined by the payment method in question (FFS or CAP). Under FFS, participants receive a fee for each medical procedure or service they provide to a patient. Under CAP, they receive a fixed payment for each patient they treat, independent of the quantity of medical services they provide. We extend Hennig-Schmidt et al.'s (2011) between-subject design by confronting doctors and medical students sequentially with both FFS and CAP payment schemes while randomizing whether participants are faced with FFS before CAP (condition FC) or CAP before FFS (condition CF). Each participant in our experiment is assigned a physician's role and joins the experiment only once, either in condition CF or in condition FC. A real patient's health is affected by the participants' treatment decisions.

Our first research question as well as our main contribution to the literature relates to bringing the field to the lab by having real doctors taking part in our lab experiment, i.e. by involving participants from the field. Our paper, thus, makes an important contribution to the literature by enhancing the external validity of lab experiments in health economics. Harrison and List (2004) highlight the importance of extending experiments beyond the standard student subject pool in order to increase their relevance for predicting field behavior. We address external validity in the present paper by investigating how practicing doctors and prospective physicians (N = 99 and 179, respectively), who are members of two non-standard subject pools (see Harrison and List, 2004), respond to FFS and CAP in a medically framed lab experiment. We are among the very first to conduct a laboratory study on payment incentives with real doctors; the other three we know of are Brosig-Koch et al. (2016, 2017b) and Fink and Kairies-Schwarz (2017).

Our second research question is concerned with the sequence effect of payment schemes and whether the order of payment systems matters when predicting behavior, as suggested, for instance, by Camerer (2003) and Harrison et al. (2005). Our within-subject design enables us to analyze this question due to the experimental feature that subjects are sequentially confronted with two systems-either FFS followed by CAP or vice versa. We are the first to study this research question in a health economics lab setting with medical doctors, as, differing from Brosig-Koch at al. (2017b), we compare both payment scheme orders.

Our third research question and contribution to the literature relates to the possibility that the optimal payment scheme is dependent on (cultural) context: The preferences of the providers are likely to be specific to the system in which they operate and are educated. Furthermore, even though financial incentives are important for physician behavior (Scott et al., 2013), other factors like training, education, professional ethics, altruism, and norms are of crucial importance as well (Yip et al., 2010). Hence, theories on the effects of incentives should ideally be confronted with empirical evidence from different cultural contextual settings. Countries like Germany and China are examples in this respect, and by conducting our experiment in China, we may test the conjecture of system-context dependence. We compare the behavioral data for medical students from China and Germany, the latter being taken from Hennig-Schmidt et al. (2011) who employed an identical design and procedure. If we find no evidence that behavior differs between the two countries; this would suggest that incentives work in similar ways despite different cultural contexts.

Our controlled experimental study in China enables additional important insights for policy makers, as the data are not confounded by, for instance, changes in patient population. Moreover, the mixture of different payment systems makes field studies rather incomparable, and identifying sequence effects on behavior will rarely be feasible with field data. In most market settings in the field, only data from one specific sequence of payment schemes will exist. In line with previous empirical and experimental evidence, we also find that in our within-subject design, both doctors and medical students provide fewer medical services under CAP than they do under FFS. Whether CAP or FFS is beneficial for the patient depends on the patient type. Patients who need many services in order to achieve maximum patient benefit are underserved under CAP compared to FFS, whereas in most cases the reverse holds for patients in need of low or intermediate levels of health care services.

Concerning external validity—our first research goal—we find small but significant differences in treatment quantity and patient benefit between medical doctors and medical students. Medical doctors appear to provide less patient benefit than medical students, and have a lower probability of choosing the patient-benefit maximizing alternative. However, this does not hold for all patients. Regarding our second research question whether the order of a payment schemes has an impact on physician behavior, we find this to hold for decisions made under the CAP scheme: Patient benefit under the CAP scheme is higher in the case where CAP is preceded by FFS than when CAP is followed by FFS. For decisions made under FFS, there is no evidence that the order of payment schemes influences patient benefit. This finding suggests that anticipating the effect of a payment reform on patient benefits requires that one consider the status quo. Concerning the dependence of behavior on cultural context, we find no significant difference when comparing Chinese and German medical students' choices under the same payment system.

The remainder of the paper is organized as follows: in Section 2, we give a brief literature account on field evidence from China and provide mixed evidence from the literature on the sequence effect. In section 3 we describe the experimental design, parameters, and procedure and state our research questions. Section 4 presents the results; and Section 5 discusses our findings and concludes.

2 Related literature on field evidence from China and the sequence effect

2.1 Field evidence from China

Countries in a transition process experimenting with introducing an appropriate physician payment system are particularly in need of information that would enable them to calibrate payment systems in an optimal way. In China, for instance, the rapidly increasing health care expenditures have become a problem in providing health care for 1.3 billion people (Eggleston, 2012). The rapid escalation in health care expenditure in China has been attributed to the fee-for-service payment systems in China[1] (e.g. Hsiao, 1995; Eggleston and Yip, 2004; Chen, 2007; Yip and Hsiao, 2008; Ramesh and Wu, 2009; Yip et al. 2010). To remedy the situation, a health care reform was introduced by the Chinese government (CPC Central Committee and the State Council, 2009; The State Council,

[1]Moreover, FFS is combined with a pricing schedule that allows China's health care providers to earn a substantial mark-up of the retail price of pharmaceuticals on each prescription (Eggleston, 2014; Yip et al, 2010).

2009). As payment reform was a priority, prospective payment methods including capitation were introduced by the Ministry of Human Resources and Social Security (MoHRSS) and the National Health and Family Planning Commission (NHFPC) in 2011 and 2012, respectively (MoHRSS, 2011, 2013). China is one of the few countries that encourages large-scale field experimentation in health to provide behavioral data on the response to changes in physician payment schemes (Heilmann, 2008). This led to a number of local 'quasi-experiments' (Eggleston et al., 2008, 150) on health-related issues, including payment reforms for physicians. The resulting evidence is mixed, possibly due to the fact that quite different combined payment schemes were introduced. Yip et al. (2014) found that capitation payment with pay-for-performance helped to reduce the prescribing of antibiotics and slightly reduced spending with no effect on other outcomes. Wang et al. (2011) conducted an experiment in rural China, changing the fee-for-service scheme of paying village doctors to a mixed payment scheme including a salary plus a bonus based on performance. Total health care spending did not significantly decrease, however, obviously due to cost shifting. A review by Yip et al. (2010) on earlier experiments suggests a potential health care cost reduction by changing provider incentives towards prospective payments. Gao et al. (2014) analyze survey data on a payment reform from fee-for-service to capitation accompanied by open enrollment in Changde, Hunan Province. The authors find that the reform did not reduce overall inpatient expenditure yet decreased patients' length of stay without reducing the quality of care.

2.2 Sequence effect

Our within-subject design enables us to analyze the behavioral effects of payment scheme sequences. Our design has the feature that payment systems are implemented in different sequences. In condition CF, participants are first confronted with CAP and then, in Part 2 of the experiment, experience. FFS. In condition FC, the sequence of payment schemes is reversed. It has been argued that the sequence of payment systems might affect behavior, but there is no clear evidence in the experimental literature on whether or not behavior is influenced by sequence.

Sequence effects were absent in a series of experimental studies on public good games. Whether subjects face a condition without (counter) punishment opportunity followed by a punishment condition or whether they are confronted with the reverse sequence had no significant effect on their behavior (see e.g. Fehr and Gächter 2000, 2002; Herrmann et al., 2008; Nikiforakis, 2008; Fischbacher and Gächter, 2010). Other examples of absent sequence effects are, for instance, experiments on the public provision of a private good (Buckley et al., 2015, 2016), on the impact of liability changes on risk taking (Füllbrunn and Neugebauer, 2013) or on the impact of moving from a pure FFS to a hybrid payment system combining FFS and pay-for-performance incentives and vice versa (Keser et al., 2014).

Sequence effects have been found in other quite different experimental settings. In experiments on the behavioral effects of introducing or removing a minimum wage, reservation wages after minimum wage abolition fall but are higher than before their introduction (Falk et al., 2006; Owens and Kagel, 2010). Hu et al., (2015) find that in an efficiency wage environment, the adjustment towards equilibrium is different for a high-price compared to a low-price condition when compared to the reverse order. In the context of myopic loss aversion, Hilgers and Wibral (2014) report that the order of confronting subjects with a change in framing a decision problem affects how they bracket their choices. Stoddard (2015) shows that behavior differs in voluntarily providing a public good when there is probabilistic uncertainty about the monetary return from the production of the public good. A negative impact of uncertainty on provision occurs only when subjects experience the condition under certainty prior to experiencing an uncertainty condition. Neugebauer and Perote (2008) observe a sequence effect in bidding behavior in first price auctions with and without information feedback. When subjects are exposed to the no-feedback condition before or after the feedback, bids in the latter are significantly higher than in the former. Heinemann and Kocher (2013) confront experimental participants with a proportionate tax regime followed by a progressive one and vice versa. They found that a regime change from a progressive to a proportionate tax tends to increase tax compliance as compared to a change in the reverse direction.

3 Experiment

3.1 Experimental design

Basic setup and decision situation

Our experimental design draws on the seminal model by Ellis and McGuire (1986). The physician is assumed to be concerned about her own profit π as well as about the patient benefit B, the latter depending on the quantity of medical services q. The specifics of the experimental design are taken from Hennig-Schmidt et al. (2011). Our experiment differs from theirs, however, in that we apply a within-subject design whereas Hennig-Schmidt et al. employ a between-subject setup.

Each participant in our experiment acts in the role of the physician. The decision task is to choose a quantity of medical services for a given patient whose health benefit is determined by that choice. Each physician i decides on the quantity of medical services $q \in 0, 1, ..., 10$ for three patient types ($j = 1, 2, 3$) with five abstract illnesses ($k = A, B, C, D, E$). She thus makes 15 decisions in both the FFS and the CAP part of the experiment. The three types of patients account for a heterogeneous patient population. Patient types reflect the patients' different states of health. The combination of patient type and illness characterizes a specific patient $1A, 1B, 1C, ..., 3D, 3E$. Patient types differ in the health benefit they gain from the medical services ($B_{1k}(q)$, $B_{2k}(q)$, $B_{3k}(q)$). The patient health benefit is measured in monetary terms. A physician's choice of medical services simultaneously determines the patient benefit and her own profit ($\pi_{jk}(q)$). The patient is assumed to be passive and fully insured, accepting each level of medical service provided by the physician. In our experiment no real patients are present. However, outside the lab, physicians' quantity choices have consequences for a real patient. The money corresponding to patient benefits aggregated over all decisions was transferred to a real patient's in-hospital account (see the Instructions in Appendix B). Thus, participants in our experiment did have an incentive to take the patient benefit into account when making their decisions.[2]

To illustrate the physicians' task, Figure 1a provides the decision screen for patient 1C under CAP whereas Figure 1b shows the decision screen for the same patient under FFS.[3] The physician gets information on her remuneration, costs and profit as well as on the patient's benefit for each quantity from 0 to 10. All monetary amounts are in Token, our experimental currency, the exchange rate being 10 Token = 1 RMB for students and 10 Token = 6 RMB for doctors (1 RMB was approximately € 0.12 at the time of the experiment).

The first two columns of the screens state the medical services and the corresponding quantities. Column 3 indicates the physician's remuneration that corresponds to a lump-sum payment per patient in CAP (Figure 1a), whereas under FFS, the remuneration increases in the quantity of medical services (Figure 1b). Column 4 shows the costs of medical services that are constant across patient types in both parts of the experiment. Physician's profit (remuneration minus costs) is given in the fifth column, and the final column comprises the patient benefit.

[2]We did not inform the participants about the name of the person to whom the money was transferred.
[3]See also the Chinese decision screens in Appendix D.

Figure 1a. Illustration of the decision screen for patient 1C under CAP

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	12.00	0.00	12.00	0.00
Service C1	1	12.00	0.10	11.90	0.75
Service C1, Service C2	2	12.00	0.40	11.60	1.50
Service C1, Service C2, Service C3	3	12.00	0.90	11.10	2.00
Service C1, Service C2, Service C3, Service C4	4	12.00	1.60	10.40	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	12.00	2.50	9.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	12.00	3.60	8.40	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.00	4.90	7.10	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	12.00	6.40	5.60	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	12.00	8.10	3.90	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	12.00	10.0	2.00	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

Parameters

Physicians are paid a lump sum of 12 Token per patient under CAP. Under FFS, physicians' remuneration increases in q. Remuneration differs with illnesses, $R_{jA}(q), R_{jB}(q), ..., R_{jE}(q)$. The lump sum paid under CAP is close to the average maximum profit per patient a subject could achieve under FFS. For an overview of all payment parameters, see panel I in Table A1 in Appendix A. The patient benefit $B_{jk}(q)$ varies across patient types. A concave benefit function is applied, the common characteristic of which is a global optimum on the quantity interval [0, 10]. There is a unique quantity q^*_{jk} that yields the highest benefit to patients of type j for illnesses k. The quantities that maximize patient benefit are $q^*_{1k} = 5$, $q^*_{2k} = 3$ and $q^*_{3k} = 7$ for patient types 1, 2, and 3, respectively—and the participants are informed of these values before they make their quantity decision. Patient benefit $B_{jk}(q)$ is shown in panel IV of Table A1. We refer to quantities smaller than q^*_{jk} as underprovision of medical care, while provision of quantities larger than q^*_{jk} is defined as overprovision. Further parameters relevant for physicians' decisions are costs $c_{jk}(q)$ and, particularly, profit $\pi_{jk}(q)$; see panels II and III of Table A.1. Physicians have to bear costs $c_{jk}(q) = 1/10 \times q^2$ under both payment systems. Under CAP, profits are the same for all illnesses. The profit-maximizing quantity $\hat{q}$ is 0 for all patients, jk. Under FFS, profits vary across illnesses because remuneration differs while costs are kept constant. The

Figure 1b. Illustration of the decision screen for patient 1C under FFS

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	0.00	0.00	0.00	0.00
Service C1	1	1.80	0.10	1.70	0.75
Service C1, Service C2	2	3.60	0.40	3.20	1.50
Service C1, Service C2, Service C3	3	5.40	0.90	4.50	2.00
Service C1, Service C2, Service C3, Service C4	4	7.20	1.60	5.60	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	9.00	2.50	6.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	10.80	3.60	7.20	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.60	4.90	7.70	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	14.40	6.40	8.00	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	16.20	8.10	8.10	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	18.30	10.0	8.30	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

profit-maximizing quantity $\hat{q}$ is 10 for all patients, jk, except for those with illness A, (i.e., patients 1A, 2A and 3A) as $\hat{q}_{jA} = 5$. For patient 1A, $\hat{q} = q^* = 5$.

3.2 Experimental protocol

Applying a within-subject design, each of the 178 Chinese medical students and 99 doctors participating in our experiment was sequentially confronted with the same 15 decisions (patients) in both of the two payment systems FFS and CAP. The subjects were randomly assigned to experimental sessions where either CAP was implemented in Part 1 of the session followed by FFS in Part 2 (condition CF) or in reversed order (FFS in Part 1 followed by CAP in Part 2, condition FC). This 2 x 2 design allows us to compare the behavior of the two subject pools over experimental conditions. Each participant was assigned a physician's role and joined the experiment only once, either in condition CF or in condition FC. Participants were informed at the beginning that the experiment consisted of two parts, but they did not know what the second part would be.

Our experiment was conducted in September 2012 and 2013 at the Center for Health Economic Experiments and Public Policy at Shandong University in Jinan, China and was programmed with z-Tree (Fischbacher, 2007).[4] Medical students who voluntarily participated in the experiment were recruited via notices posted at the campus and by email invitations. Doctors were recruited through a phone call stating that a research experiment from Shandong University needed volunteers.[5]

The experimental procedure was as follows and was exactly the same for medical students and doctors. After having arrived and before the experiment started, participants were randomly allocated to their workstations. The workstations were numbered and separated from each other by wooden panels and curtains. It was thus guaranteed that they made their decisions in both parts of the experiment in complete anonymity. Then, instructions for Part 1 of the experiment were distributed to participants and read out by a native experimenter. Participants decided under either a CAP or an FFS system. Subjects were given plenty of time to read the instructions and to ask clarifying questions in private, which were answered individually. In cases that the content was important for all participants, the question and answer were repeated in public. To check for participants' understanding of the decision task, they had to answer a set of test questions on remuneration, costs, physician profit and patient benefit for three different quantities of medical services for a patient they were not confronted with in the actual experiment.[6] Each participant then went through a sequence of 15 choices (patients) on the quantity of medical services to be provided. The order of patients to be treated was predetermined and kept constant across conditions. After each decision, each participant in both parts of the experiment was informed about his/her profit and the patient benefit generated by the previous choice. At the end of the first part of the experiment, each participant received information about his/her total profit achieved and the total health benefit generated during all 15 quantity decisions. Finally, the participants answered some open-ended questions.

Next, instructions for the second part of the experiment were distributed and read out by the native experimenter. In Part 2, participants decided under the payment system they had not yet been confronted with. Again, each decision maker received information on his/her total profit achieved and the total health benefit created during all 15 decisions. After the second part of the experiment had been completed, participants were again asked some open-ended questions. The doctors were also asked about socio-demographic variables and professional experience. Finally, participants were informed about their individual total profit and the resulting total benefit aggregated over Parts 1 and 2 of the experiment

[4]All material distributed to the participants was translated into Chinese by a Chinese native fluent in German from the original German version by using the back translation method (Brislin, 1970). For a translation into English, see Appendix C1. It is important to instruct participants in their native language because the language the experiment is presented in may affect their behavior; see e.g. Costa et al. (2014a,b).

[5]When signing up, participants did not know about the decision tasks.

[6]See Appendix C2 for the English translation of test questions and the respective computer screens.

as well as on their final monetary payoff. Finally, participants were paid in private and dismissed individually.

To ensure that the experimental subjects trusted the experimenters to actually transfer the money derived from the patient benefit, a certain procedure was applied to ensure trust[7]: A monitor was randomly selected from the participants in a session. He/she verified the amount of money corresponding to the patient benefits aggregated over all decisions of all participants in the respective session. Then, the monitor and an assistant to the experimenters went by taxi to the Shandong Cancer Hospital in Jinan and paid the corresponding amount in cash into the patient's account at the hospital's cashier's desk.[8] The monitor signed a statement that the appropriate monetary amount was paid into the patient's hospital account. All participants in each session received an email stating the amount equaling the aggregate health benefits generated during the respective session. Each monitor in the medical student subject pool was paid an additional 50 RMB and each doctor 200 RMB.

We conducted four sessions, and each experimental session comprised one condition with conditions alternating across sessions. Sessions lasted for about 90 minutes. Based on the decisions in the two conditions, each of the 178 medical students on average earned 28 RMB; 15 RMB (€1.80) in CAP and 13 RMB (€1.56) in FFS plus a show-up fee of 15 RMB (€1.80). Doctors on average earned 160 RMB (86 RMB (€10.32) in CAP and 74 RMB (€8.88) in FFS.[9] Based on all 8,310 decisions, a total of 19,814 RMB (€2,377.68) was transferred to the real patient's account; 4,751 RMB (€570.12) for the sessions with medical students and 15,063 RMB (€1,807,56) for the sessions with doctors.

3.3 Research questions

Our first research question is: *do medical students behave differently from medical doctors?* The question relates to the external validity of experimental studies where medical students are assigned the role of medical doctors. Harrison and List (2004) highlight the importance of extending experiments beyond the standard student subject pool to increase their relevance for predicting behavior in the field. We address external validity in the present paper by bringing the field to the lab (Cardenas 2005) and having practicing doctors

[7]This procedure is similar to Eckel and Grossman (1996), Hennig-Schmidt et al. (2011), Godager and Wiesen (2013), Hennig-Schmidt and Wiesen (2014), Godager et al. (2016) and Brosig-Koch et al. (2016, 2017a).

[8]We took great care that the monitor did not see the name of the real patient in order to maintain the patient's anonymity.

[9]Average payoffs for students approximately corresponds to the hourly wage of a student helper at Shandong University of about 30 RMB. For doctors the average hourly wage is about 120 RMB.

participate in our experiment. We also involve prospective physicians, i.e. medical students, another non-standard subject pool (Harrison and List, 2004). Moreover, our sample sizes (99 doctors and 178 medical students) are much larger than in previous comparable experimental studies (see e.g. Hennig-Schmidt et al., 2011; Hennig-Schmidt and Wiesen, 2014; Brosig-Koch et al., 2016, 2017a; Cox et al., 2016b).

We systematically analyze how the members of these two non-standard subject pools respond to the financial incentives of CAP and FFS and if so whether doctors and medical students behave differently. Differences would suggest that professional experience in treating patients change physicians' provision behavior—to the benefit or the detriment of the patients. Brosig-Koch et al. (2016), for instance, find that doctors and medical students respond to the incentives in a consistent way: significantly more medical services are provided under FFS than under CAP. The intensity, however, by which subjects respond, differs by subject pool. Physicians are more patient-regarding than medical students in both payment schemes. The first research question also relates to whether professional experience matters for decisions insofar as doctors and medical students behave differently, in particular with regard to providing the maximal patient benefit.

The second research question is: *does the sequence in which participants experience the payment systems influence their behavior?* Is, for example, behavior under a CAP scheme different if CAP follows FFS compared to the case where CAP is introduced first?

Our third research question is: *is provision behavior dependent on cultural context?* The importance of non-financial factors like culturally different education and norms for physician behavior, has been highlighted in recent literature (Yip et al., 2010; Scott et al., 2013). To test the hypothesis of cultural context dependence, we compare behavioral data for medical students from China and Germany based on experiments with identical parameters. We have no clear conjecture as to possible behavioral effects. However, finding no evidence for behavioral differences between the two countries would suggest that incentives work in similar ways despite different cultural contexts.

4 Results

In subsection 4.1, we give a brief description of the subject pools and test for differences in aggregate provision behavior between CAP and FFS. In subsection 4.2, we analyze and test whether doctors and medical students are different in providing treatment and patient benefit. In subsection 4.3, we analyze how the sequence of payment schemes affects behavior and report regression results where we jointly assess the effect of payment schemes, payment scheme sequences and subject pools on students' and doctors' behavior. Subsection

4.4 compares behavior in the two subject pools of medical students from Germany and China. Throughout the paper, all statistical tests applied are two-sided.

4.1 Subject pool characteristics

In our experiment, 277 subjects participated (Table 1). Of these, 178 were medical students of whom 57 % were females. The overall average duration of study was 4.9 semesters. The major of all medical students was Clinical Medicine. The number of participating doctors was 99 with an average age of 40, and 70 % were females. They had on average of 16.23 years of professional experience (min. 1, max. 39). The doctors were practicing as general practitioners (75 %), in traditional Chinese medicine (10 %) or in public health (4 %); 11 % of the doctors practiced in all or several of these fields. All doctors were employed at community health centers where salaries are set according to a fixed salary scheme. Thus, both the medical students and the doctors have in common that they had little or no practical experience with fee-for-service payment or capitation payment systems.

Table 1. Experimental conditions, number of subjects, decisions and participants' characteristics

	Subject pool			Total #	Total #
Condition	Doctors	Students	male	subjects	obs
CF: CAP followed by FFS	49	88	36%	137	4110
FC: FFS followed by CAP	50	90	24%	140	4200
Total	99	178	30%	277	8310

Notes:
This table shows experimental conditions, number of subjects and their decisions as well as participants' characteristics. Total #obs is the total number of decisions per condition

The aggregate provision behavior under CAP and FFS is presented in Table 2. Here we analyze the data pooled over doctor and medical student subject pools (N=277 subjects; 4155 decisions per payment system). We here also pool data from the same payment-scheme, regardless of whether the scheme was implemented first or second in the experiment. In line with earlier studies, we find that our participants respond to the incentives given by the payment systems: average quantities in CAP are lower than those in FFS (CAP: 4.56, FFS: 6.10; N=277). This also holds true for patient benefit but only slightly (CAP: 8.62, FFS: 8.70); see Table 2 for both q and B. Analyzing the mean quantity provision for each of the 15 patients in CAP and FFS, we observe a significantly higher number of services under FFS than under CAP ($p \leq 0.0001$, Wilcoxon matched-pairs signed-ranks test, WM in the following)[10]. Thus, we corroborate the experimental results of Hennig-Schmidt et al.

[10] Applying a conservative Bonferroni correction for multiple hypothesis testing gives an adjusted threshold for statistical significance of $p = 0.05/15 = 0.0033$ when tests are applied 15 times. In the above analysis, the difference for patient 1A becomes insignificant. In the following, we will explicitly mention whenever the use of the Bonferroni correction would imply not rejecting the null hypothesis.

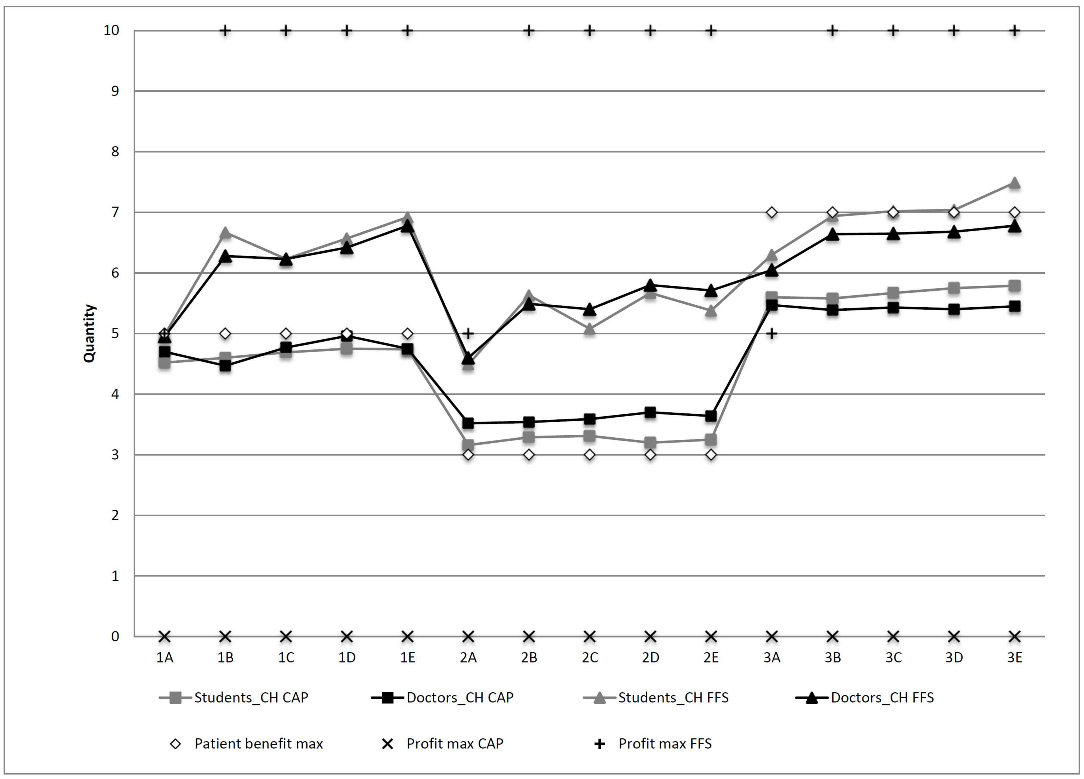
(2011), Keser et al. (2013), Hennig-Schmidt and Wiesen (2014) and Brosig et al. (2016, 2017a) showing that under FFS more services are provided as compared to CAP. We find the same behavior in our within-subject design.

Table 2. Aggregate provision behavior under CAP and FFS. Mean (Std.Dev) of quantity and patient benefit, and the number of decisions

Payment	Doctors			Medical students			Total		
system	Quantity	Benefit	#obs	Quantity	Benefit	#obs	Quantity	Benefit	#obs
CAP	4.59 (1.78)	8.29 (2.58)	1485	4.53 (1.57)	8.81 (2.17)	2670	4.54 (1.65)	8.62 (2.19)	4155
FFS	6.03 (1.92)	8.61 (1.75)	1485	6.16 (1.78)	8.98 (1.19)	2670	6.11 (1.83)	8.70 (1.82)	4155

Notes:
This table shows descriptive statistics on quantities of service provision, provision of patient benefit over payment systems and subject pools. #obs is the number of decisions under each payment scheme aggregated over conditions CF and FC.

Figure 2. Mean quantity provision for each of the 15 Patients under CAP and FFS differentiated according to subject pools – pooled over both parts of the experiment.



Notes: This figure shows average quantities of service provision as well as patient benefit and profit maxima for payment systems FFS and CAP for Chinese doctors (N=99), and medical students (N=178), pooled over both parts of the experiment.

We find that under CAP, the chosen quantities are significantly different from q^* for all 15 patients (p $\leq$ 0.0002, WM). Underprovision prevails more often than overprovision for all patients of patient types 1 and 3, while for patient type 2 the reverse holds true. Under FFS chosen quantities deviate significantly from q^* for 12 of the 15 patients (p $\leq$ 0.0337, WM)[11]. The three patients where we do not find significant differences are patients 1A—where $q^*$and $\hat{q}$ coincide—3C, and 3D. Overprovision is more common than underprovision except for patients of type 3.

Summary: In line with previous empirical and experimental studies, the incentives of the two payment systems affect medical service provision in that participants provide more services under FFS than under CAP. Over- and underprovision in the two payment systems depend on patient type.

4.2 External validity—Do Chinese doctors and medical students differ in their provision behavior?

Our first research interest is concerned with adding to the external validity of laboratory experiments on physician payment incentives by bringing the field to the lab and having real doctors in addition to prospective physicians take part in our experiment. To this end, we compare how Chinese doctors and medical students respond to FFS and CAP in providing services q and in the resulting patient benefit B. In this subsection, we again analyze the data pooled over both parts of the experiment for the same payment system.

Our analysis shows that doctors and medical students are influenced by the incentives of the payment systems. Doctors provide an average of 4.59 (s.d. 1.78) services in CAP and 6.03 (s.d. 1.92) services on average in FFS, see Table 2. Mean quantities provided by students are 4.53 (s.d. 1.57) services on average in CAP and an average of 6.16 (s.d. 1.78) services in FFS. Our separate analysis of mean quantity provision for each of the 15 patients in CAP and FFS corroborates for both subject pools the significantly higher service provision under FFS than under CAP (doctors: $p \leq 0.0133$, medical students: $p \leq 0.0001$, WM)[12] found in subsection 4.1 (see also Figure 2). Doctors and medical students differ in how strongly they respond to the incentives in both payment schemes. Doctors provide significantly less patient-regarding services for each of the three patient types ($p \leq 0.0366$ Mann-Whitney U-test, MWU in the following) except for patient type 1 in CAP; see also Fig. 2.[13] This is in contrast to results of Brosig-Koch et al. (2016) from Germany who report that doctors are influenced by both payment schemes but less so than medical students, i.e. they provide more patient-regarding treatment. We will come back to this issue in the conclusion.

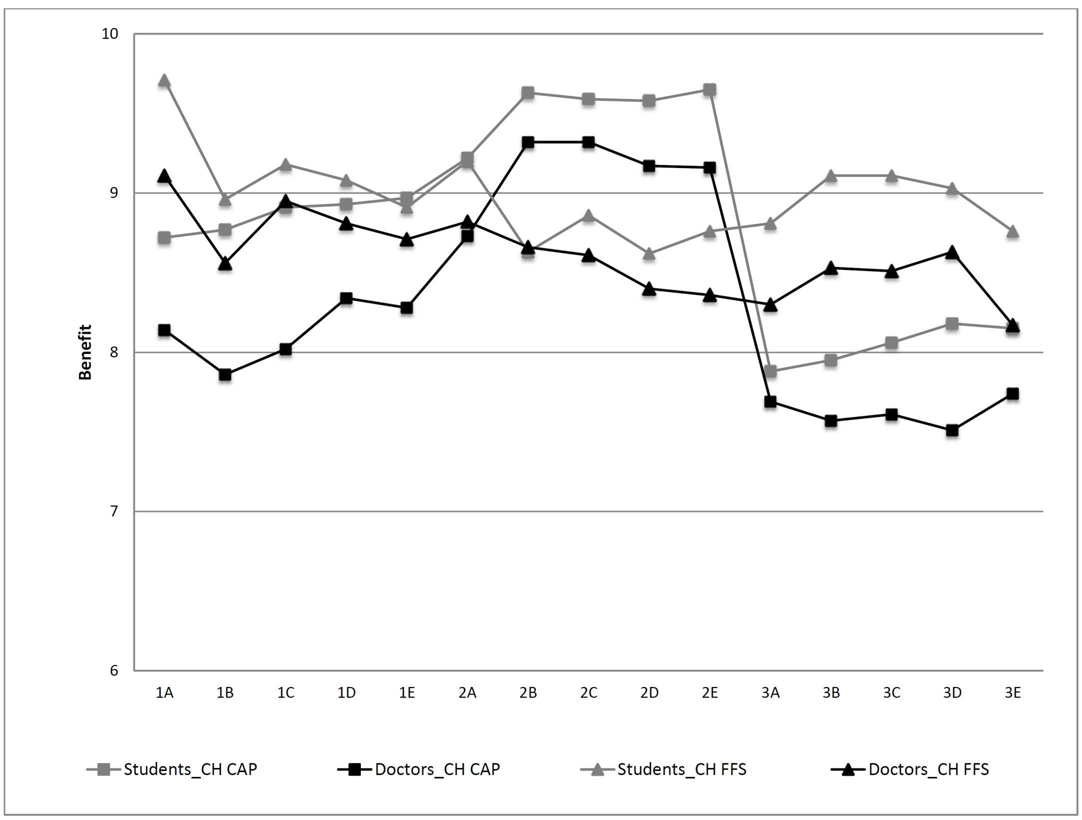
[11] For decision 3B (p=0.0337), the Bonferroni-corrected p-value does not allow for rejecting the hypothesis of no deviation from q^*.

[12] For decision 1A ($p = 0.0133$), the Bonferroni-corrected p-value does not allow for the rejection of the hypothesis of no difference between CAP and FFS.

[13] The same holds true for patient benefit for all patient types in both payment schemes ($p < 0.0001$, MWU), see also Figure 3.

To get deeper insights into the causes of these differences, we next compare whether doctors and medical students behave differently when providing medical services q and patient benefit B for each of the 15 patients. We find that under FFS, doctors' choices result in a significantly lower q for four of the patients of Type 3 (3B–3E)[14]. With regard to patient benefit, doctors provide significantly lower patient benefit for patients 1A and 3A–3E. In CAP, doctors' service provision q is significantly lower for all patients of Type 2 and slightly significantly higher for patient 3B[15]. As for B, doctors provided significantly lower patient benefit for all patients of Type 1, for 2A-2D and 3D. Note that we do not find any patient for whom physicians provide significantly more patient benefit than medical students. For p-values of the 15×4 MWU-tests see Table B1 in Appendix B.

Figure 3. Mean patient benefit for each of the 15 Patients under CAP and FFS differentiated according to subject pools – pooled over both parts of the experiment.



Notes: This figure shows provision of patient benefit for payment systems FFS and CAP for Chinese doctors (N=99), and medical students (N=178), pooled over both parts of the experiment.

The differences between doctors' and medical students' behavior might be due to the

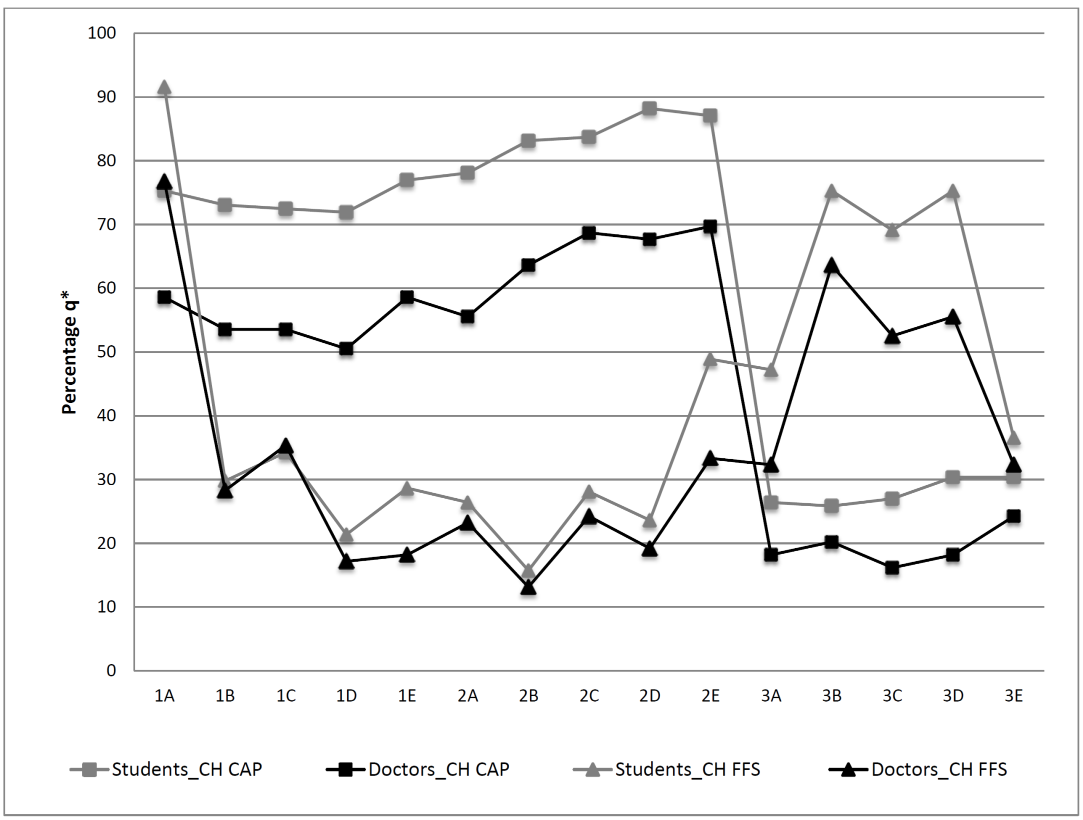
[14] With a Bonferroni correction, we no longer reject the hypothesis that a difference exists for the treatment of patients 3B and 3C.

[15] When applying a Bonferroni correction, the differences are only significant for patient 2D.

fact that the former less often provide B^{max}, the patient-optimal quantity q^*. Figure 4 shows the relative frequencies of subjects choosing to maximize patient benefit. We observe substantial differences with regard to subject pools, patients, and payment systems. Pooled over both payment schemes, 52.7 % of medical students chose q^* compared to only 40.7 % of doctors. This difference is highly significant (p < 0.001, chi^2 test). Testing behavioral differences for each patient and payment system reveals that under CAP, the doctors are significantly less likely to provide q^* (B^{max}) for 80 % of the patients (12 of 15) (p ≤ 0.05 chi^2 tests); the exceptions are patients 3A, and 3B, 3E.[16] Under FFS, this applies to 33 % (5 of 15), namely 1A, 2E, and 3A–3C (p ≤ 0.05)[17].

Result 1: From our non-parametric analysis, we find that both doctors' and medical students' service provision is higher in FFS than in CAP with doctors providing the patient optimal quantity less often in both payment schemes. Doctors' decisions lead to lower patient benefit than those of medical students. This holds more strongly for CAP than for FFS.

Figure 4. Percentage of subjects who choose q^* for the 15 decisions in CAP and FFS.



Notes: This Figure shows the proportion of subjects providing q^* differentiated by payment systems and subject pools. Data are pooled over both parts of the experiment. N=178 for medical students and N=99 for doctors.

[16]When applying a Bonferroni correction, the difference for patient 3C in CAP becomes insignificant.
[17]When applying a Bonferroni correction, the difference for patient 3B in FFS becomes insignificant.

Table3. Aggregate quantities and patient benefits by condition and subject pool.

	Country (Subject type)	Condition	N	# obs	Part 1 of the experiment Payment	Mean	SD	Part 2 of the experiment Payment	Mean	SD
q	**China**									
	(All)	CF	137	4110	CAP	4.57	1.8	FFS	6.06	1.83
		FC	140	4200	FFS	6.16	1.83	CAP	4.53	1.49
	(Doctor)	CF	49	735	CAP	4.56	1.9	FFS	5.92	1.89
		FC	50	750	FFS	6.14	1.95	CAP	4.61	1.65
	(Student)	CF	88	1320	CAP	4.57	1.75	FFS	6.14	1.8
		FC	90	1350	FFS	6.18	1.76	CAP	4.48	1.38
	Germany									
	(Student)	N/A	22	330	CAP	4.4	1.64	N/A		
		N/A	20	300	FFS	6.6	1.85	N/A		
$B(q)$	**China**									
	(All)	CF	137	4110	CAP	8.28	2.66	FFS	8.91	1.41
		FC	140	4200	FFS	8.79	1.45	CAP	8.97	1.91
	(Doctor)	CF	49	735	CAP	7.89	2.86	FFS	8.58	1.89
		FC	50	750	FFS	8.63	1.62	CAP	8.7	2.2
	(Student)	CF	88	1320	CAP	8.49	2.52	FFS	9.09	1.01
		FC	90	1350	FFS	8.88	1.34	CAP	9.13	1.7
	Germany									
	(Student)	N/A	22	330	CAP	8.56	2.46	N/A		
		N/A	20	300	FFS	8.83	1.1	N/A		

Notes: This table shows descriptive statistics on quantities of service provision and patient benefit for conditions, payment systems for all participants as well as for subject pools. The data are differentiated by parts 1 and 2 of the experiment. N is the number of independent observations by conditions and payment systems, whereas # obs is the number of decisions under both payment schemes in each of the conditions CF and FC.

4.3 Sequence of payment schemes.

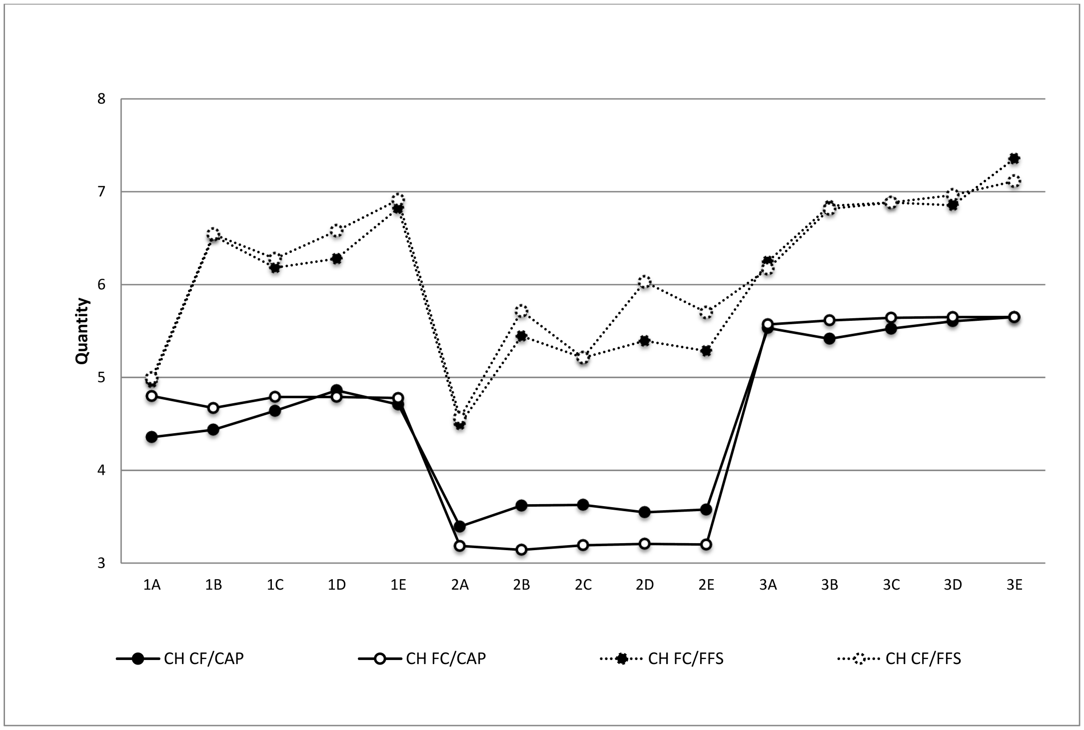
Our second research interest is concerned with the effect of payment scheme sequence. Does experiencing CAP first and FFS second or vice versa affect the behavior of doctors and medical students? Our sequential within-subject design enables us to analyze this question.

Non-parametric analysis.

Table 3 shows the average provision of services q and patient benefit B by condition and differentiated by parts 1 and 2 of the experiment for all participants and for doctors and medical students separately.

Figure 5 presents the mean quantity provision for each of the 15 Patients under CAP and FFS, pooled over the two Chinese subject pools and differentiated by parts 1 and 2 of the experiment. We see that although the pattern of mean quantity choices have similarities in both parts of the experimental sessions, participants behave (slightly) more patient regarding in part 2 meaning that the chosen q is closer to q^* for many patients. This is particularly true for CAP.

Figure 5. Mean quantity provision for each of the 15 Patients under CAP and FFS in part 1 and 2 of the experiment pooled over the Chinese subject pools.



Notes: This figure shows mean quantity provision for each of the 15 Patients for both payment systems CAP and FFS in both parts of the experiment. The data are pooled over the Chinese subject pools. N=277.

We conducted a series of tests to analyze whether a systematic tendency concerning the sequence of payment schemes exists and test this for both CAP and FFS and for each of the 15 decisions[18]. For service provision q, we find that for 7 of the 15 decisions made under CAP, the sequence does matter in that we find (weakly) statistically significant differences in q between parts 1 and 2 of the experiment. Depending on patient type, q is higher (1A, 1B) or lower (2A–2E) in FC, i.e., when CAP is implemented second[19]. We only find q to be significantly higher under FFS in CF for 2D[20]. This finding is corroborated by applying a robust test for equality of variance (Levene, 1960) where the variance of q-choices under CAP is found to be significantly larger in CF than in FC ($p < 0.0001$). In contrast, there is no evidence in the data that this holds true for FFS. Thus, we cannot reject the null hypothesis of equal variance under FFS ($p = 0.2560$). For patient benefit, we get similar findings as those for q-choices with more pronounced benefit-increasing behavior in CAP (see Figure 6). For 9 of the decisions made under CAP, the sequence does matter in that B in FC is highly significantly higher in 1A–1C, 1E, and in 2A–2E, i.e., when CAP follows

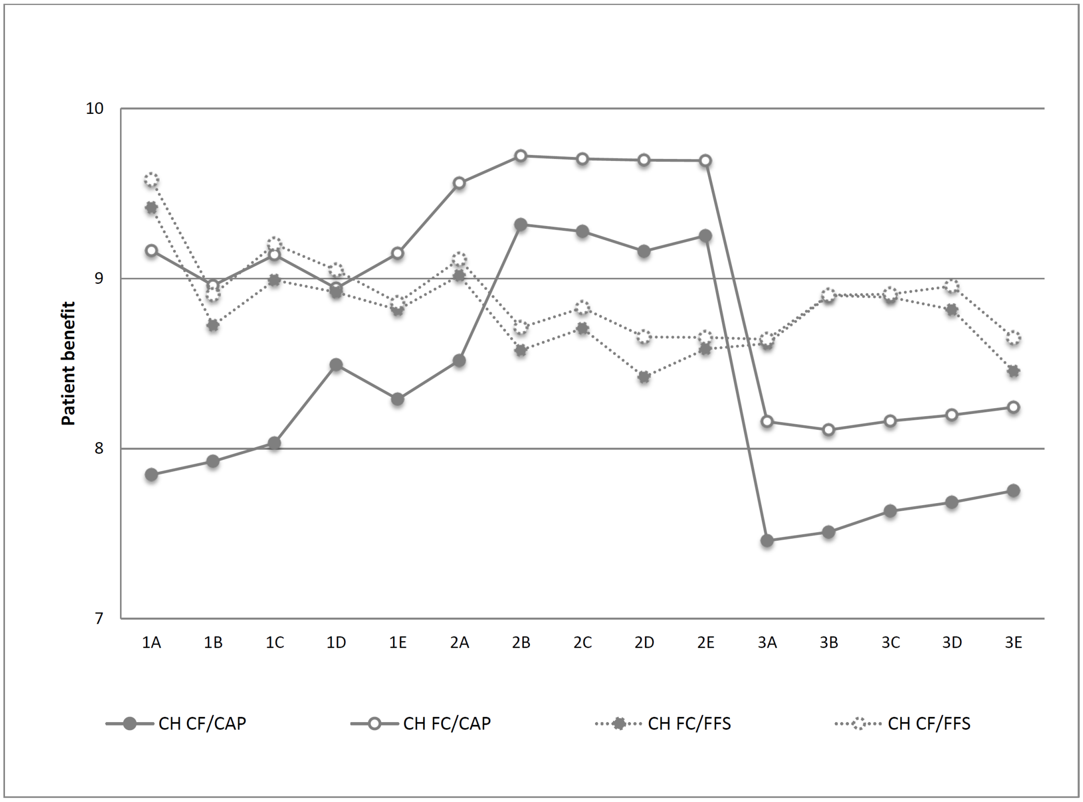
[18]We present the p-values for the 15 × 4 MWU tests for both quantity and patient benefit in Table B2 of Appendix B.

[19]When applying a Bonferroni correction, the differences are significant only for patients 2B and 2C.

[20]The difference becomes insignificant when applying a Bonferroni correction.

FFS. We only find a significant sequence effect under FFS in 2D[21].

Figure 6. Mean provision of patient benefit for each of the 15 Patients under CAP and FFS in part 1 and 2 of the experiment pooled over the Chinese subject pools.



Notes: This figure shows mean provision of patient benefit for each of the 15 Patients for both payment systems CAP and FFS in both parts of the experiment. The data are pooled over the Chinese subject pools. N=277.

Regression analysis.

In order to jointly assess the effects of different sequences in payment schemes, payment schemes in general, and differences in the behavior of doctors and medical students we finally run two regression models where we let the chosen quantity and the patient benefit correspondingly enter as dependent variables for each model. We include all 30 decisions of each of the 99 doctors and 178 medical students. By fitting random effects models and applying clustered standard errors, the regression models enable us to handle repeated observations of the same subject. Since we have two different payment schemes (CAP and FFS) and two different sequences of payment schemes (CF and FC), each decision will be made in one of four scenarios. We, therefore, create dummy variables for the three settings, FFS in CF, CAP in CF, and FFS in FC with decisions made under CAP in FC being the

[21] This becomes insignificant when applying a Bonferroni correction.

Table 4. Results from ordinal regression models. The impact of payment schemes and sequence on service provision, q and patient benefit, $B(q)$.

Regressors	Dependent variable: q Odds Ratio		95% C.I.		Dependent variable: $B(q)$ Odds Ratio		95% C.I.	
FFS in CF	6.84	***	5.35	8.74	0.42	***	0.31	0.58
CAP in CF	1.07		0.84	1.38	0.46	***	0.32	0.67
FFS in FC	7.71	***	5.74	10.35	0.35	***	0.28	0.42
Doctor	0.95		0.77	1.16	0.60	***	0.44	0.81
Patient								
1A (reference)								
1B	2.80	***	2.42	3.24	0.15	***	0.11	0.22
1C	2.44	***	2.13	2.80	0.25	***	0.18	0.35
1D	2.95	***	2.58	3.37	0.19	***	0.14	0.27
1E	3.54	***	2.96	4.22	0.19	***	0.13	0.27
2A	0.28	***	0.23	0.33	0.23	***	0.16	0.34
2B	0.52	***	0.40	0.67	0.22	***	0.15	0.32
2C	0.40	***	0.31	0.53	0.25	***	0.18	0.36
2D	0.53	***	0.41	0.69	0.20	***	0.14	0.29
2E	0.40	***	0.28	0.57	0.31	***	0.21	0.45
3A	4.66	***	3.78	5.76	0.04	***	0.03	0.05
3B	7.07	***	5.76	8.68	0.04	***	0.03	0.06
3C	7.63	***	6.19	9.42	0.05	***	0.03	0.06
3D	7.67	***	6.24	9.42	0.05	***	0.03	0.07
3E	10.67	***	8.43	13.51	0.04	***	0.03	0.05
Number of obs				8310				8310
Number of subjects				277				277
Obs per subject				30				30
Random effects				Yes				Yes
Clustered S.E.				Yes				Yes
Log pseudolikelihood				-13894.3				-16439.3

Note: The table shows the effect of payment schemes and sequence estimated by means of ordinal logistic regression with subject-specific random effects based on decisions made by medical doctors and medical students in China. Estimates with *** indicate that the odds ratio is significantly different from 1 at the 0.1% level in a two-tailed test.

reference category. Furthermore, we include a dummy for doctors. Finally, we control for patients with 14 dummy variables, where patient 1A enters as reference category. We estimate the models with individual specific random effects and present robust standard errors clustered at the level of the individual. The results are presented in Table 4.

The regression analyses corroborate the results of our non-parametric analyses. Estimated odds ratios of FFS in FC and CF are both substantially higher (lower) than 1, when quantity (benefit) is the dependent variable. The interpretation is that FFS results in larger quantities than CAP. We observe that for patient benefit the estimated odds ratio of the dummy for doctors is less than 1, implying that decisions made by medical doctors on average lead to lower patient benefit than those made by medical students. Analyzing the sequence effect of CAP on provided quantity and patient benefit, we observe that the estimated odds ratio of CAP in CF is not significantly different from 1 for quantity but is significantly less than 1 for patient benefit. The interpretation is that CAP in FC results in higher average patient benefit than CAP in CF. In both regression models, the confidence intervals of FFS in CF and in FC have substantial overlap, implying that there is no significant difference between the two conditions CF and FC for FFS with regard to patient benefit. Our findings regarding the CAP-system suggest that for anticipating the effects of a physician payment reform on provision behavior and on patient benefits in particular, it may be important to take the status quo from which the reform starts into account. We will come back to this issue in the discussion in section 5.

Result 2: We find a sequence effect in that CAP results in higher patient benefit when it is experienced after FFS as in FC than when is experienced first as in CF. We find no significant difference between the two conditions CF and FC for FFS with regard to patient benefit. Medical doctors provide less patient benefit than medical students.

4.4 Comparing behavior of medical students from China and Germany.

Our third research question concerns whether different contextual settings influence behavior. To this end, we compare behavioral data for medical students from China and Germany, the latter being taken from Hennig-Schmidt et al. (2011), who did not employ a within-subject design that confronts subjects with two different payment schemes[22]. Apart from that, our design and procedure are identical to theirs. For the comparison, we use the behavioral data from the first parts of our conditions CF and FC for Chinese medical students. We thus analyze data for N=88 Chinese and N=22 German medical students in CAP and for N=90 and N=20, respectively, in FFS (see Table 3). Figure 7 shows that Chinese and German medical students behave rather similarly. The latter consistently provide (slightly) more medical services on average in FFS than Chinese medical students do (see Table 3). We, however, find no evidence in the data that chosen quantities q or the provided patient benefit B differ significantly between the two subject pools when applying MWU-tests for each of the 15 patients under both payment systems[23]. We also test whether the frequency of choosing q^* is different between the two subject pools. Overall, nearly half of all patients are provided with their benefit maximum (German students: 49.7 %, Chinese students 47.9 %). Differences are small and not statistically significant ($p = 0.148$, Fisher's exact test). We also do not find statistically significant differences between the two student groups in the frequencies of choosing q^* for each patient within the two payment systems ($p > 0.13$, Chi^2 test). Exceptions are patients 1A in FFS and 2C in CAP, where the difference is weakly significant ($p > 0.06$, Chi^2 test) [24]. As robustness

[22]Hennig-Schmidt et al. (2011) had no doctors participating in their experiment.

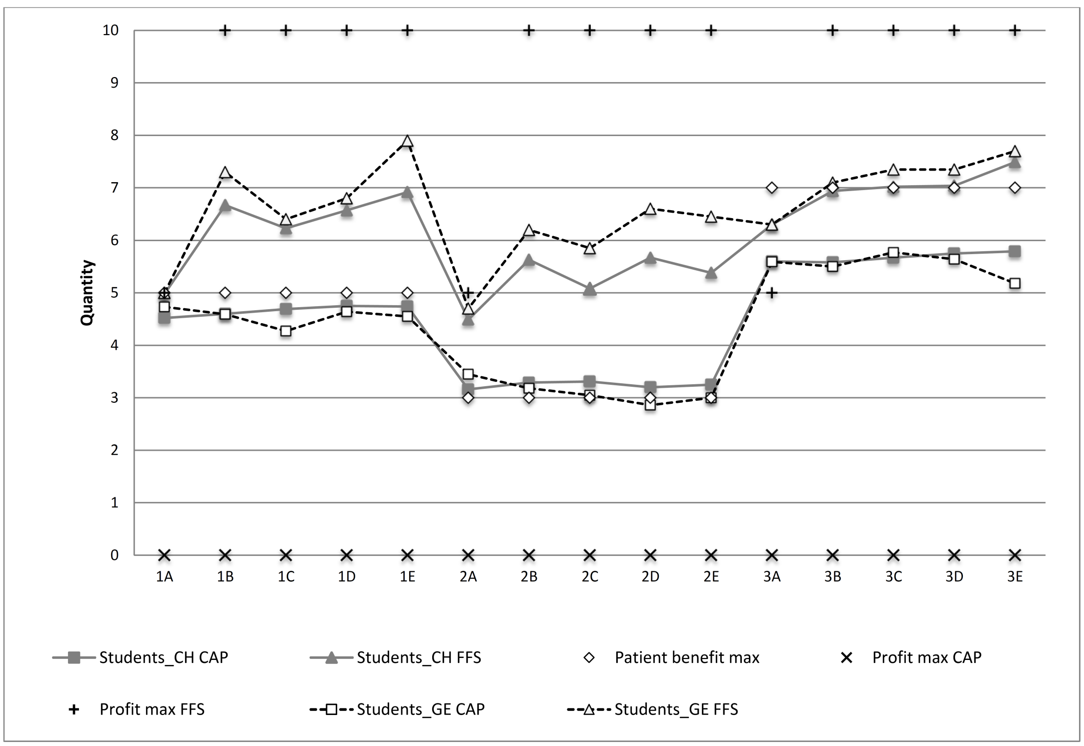
[23]See the values from the 4*15 MWUtests in Table B3 in Appendix B.

[24]See the p-values of the 2×15 chi^2 tests in table B4 in the appendix. Applying a Bonferroni correction also renders patients 1A in FFS and 2C in CAP insignificant.

check and to jointly assess whether there are differences between the three subject pools (Chinese doctors and Chinese and German medical students), we run regression models where we let the chosen quantity and the patient benefit correspondingly enter as dependent variables. Also in the regression analysis, we only include the first 15 decisions from the Chinese sessions. The previous non-parametric results of significant differences between doctors and medical students in China are confirmed as well as our findings of no significant differences between the two groups of medical students (see Table B5 in Appendix B).

Result 3: Chinese and German medical students behave similarly in providing services q and patient benefit B as well as in choosing the optimal patient treatment.

Figure 7. Mean quantity choices for each of the 15 Patients under CAP and FFS for the Chinese and the German subject pool.



Notes: This figure shows mean quantity provision for each of the 15 Patients for both payment systems CAP and FFS of Chinese and German medical students. China: medical students CAP: N=88, FFS: N=90; Germany: medical students, CAP: N=22, FFS: N=20.

5 Discussion and concluding remarks.

In this paper, we examine how changes in payment mechanisms influence the behavior of medical students and doctors. We focus on the two payment methods fee-for-service and capitation and analyze whether providers behave differently when FFS is introduced first and CAP second, as compared to when payment mechanisms are introduced in opposite order. To this end, we introduce a fully incentivized laboratory experiment and implemented controlled ceteris paribus changes of the payment schemes. A further main goal is to add external validity to experimental studies by having practicing doctors and prospective physicians, i.e. medical students, participate in our experiment. Finally, we analyze whether different cultural contextual settings matter for physicians' provision behavior. We therefore conduct our experiment in China and compared the behavioral data for medical students from China and Germany.

Our analysis shows that the incentives of the two payment systems affect medical service provision in that participants provide more services under FFS than under CAP, which is in line with previous empirical and experimental studies. Over- and underprovision in the two payment systems depend on patient type. These results hold true for doctors and medical students alike. We find that the behavior of medical doctors and medical students differ, however, in that doctors not only choose the patient optimal quantity less frequently in both payment schemes but also provide lower benefit to patients than medical students. This result is more apparent in the case of CAP than for FFS.

Another main insight is that the sequence of payment schemes affects physician provision behavior. We find that CAP results in a higher provision of patient benefit when it is experienced after FFS as compared to when it is experienced first. For FFS, we find no evidence of a sequence effect.

Finally, we find that Chinese and German medical students behave similarly in providing medical services and patient benefit as well as in choosing the patient-optimal treatment. This suggests that the financial incentives of the payment systems in our experiment work in a similar way independent of the cultural context in which our participants are educated and operate. An implication of this finding is that results from health economic laboratory experiments can provide broad knowledge on expected behavior under cultural and institutional contexts that are different from where the actual experiment is conducted.

Our finding that Chinese doctors provide less patient-regarding treatment than Chinese medical students seems to suggest that growing professional experience in treating patients may change physicians' provision behavior after graduating and entering medical practice—in our experiment to the detriment of patients. Intuitively, however, one might expect additional experience to improve performance (Epstein et al. 2013). However, empirical evidence on the subject is mixed. Reid et al. (2010) on the one hand do not find any association between experience and performance (see also the literature cited therein).

McAlister et al. (2015) found no negative association, and Epstein et al. (2013) observed a positive relationship between physician experience and patient outcomes. On the other hand, Choudhry et al. (2005) systematically review studies relating experience and age to physician performance and find that 70 % of these studies demonstrate a negative association between length of time in practice and patient outcomes as well as several other measures of good physician performance. The authors conclude that physicians with more experience may be at risk of providing lower-quality care (see also Pham et al. 2005, Weinberger et al. 2005, Streja and Rabkin 1999). Choudhry et al. (2005) recommend giving special attention to those physicians who may need quality improvement interventions.

We also find that subject pools respond to financial incentives in a similar and consistent way, and this is in line with Brosig-Koch et al. (2016)—the only published experimental study that analyzes both doctors' and medical students' reactions to different payment schemes (in Germany). However, different from our study, German doctors are more patient-regarding on average in both payment systems by providing less (more) services in FFS (CAP) as compared to medical students. One explanation for this difference is suggested by the aforementioned comment of Choudhry et al. (2005) on quality improvement interventions. Furthermore, German physicians are required to take mandatory, annual advanced education and training courses, and the doctor sample of Brosig-Koch et al. (2016) was recruited from such courses. This engagement may have made the German doctors more aware of patients' needs and thus less prone to the 'curse of experience' (Abbink and Rockenbach, 2006)—i.e. the phenomenon that practitioners apply certain patterns of behavior to the experimental situation that are adequate for their situation in professional life but are not optimal for a contextually framed but still stylized experimental setup like ours.

Our further main finding that the order in which payment schemes are introduced may matter, is in line with a substantial part of the experimental literature reviewed in section 2. In our experiment, CAP results in higher patient benefit when it is experienced after FFS in FC than when it is experienced first in CF. No such evidence was found for FFS. One might argue that a learning effect accounts for the difference found. This cannot be excluded, as subjects get more familiar with the benefit parameters of the three patient types. However, if learning were the main factor, there should be a similar effect in FFS. This is not the case, however. Looking at the data, we observe that subjects seem to be particularly hesitant to provide relatively low quantities for type 2 patients in CF when they experience CAP first, even though no tradeoff between own profit and patient benefit exists in this situation. Lower quantity choices would not only have increased benefit for B-patients but also the subject's profit (see Figure 2 and Table A1 in Appendix A). This may point to the subjects being motivated by a perception that ‘more is better’ (see, e.g., Anderson and Chalkidou, 2008) in their first encounter with a CAP system. See, for instance, Abbink and Rockenbach (2006) and Cooper et al. (1999) for similar findings for German professional traders or Chinese managers, respectively (c.f. also Frechette, 2015).

The general advice of cautiousness when deriving policy implications from experimental results apply. However, our findings corroborate the findings from the field that FFS payment encourages higher service volumes than CAP and services volumes under FFS can become higher than what is in the best interest of the patient. For CAP-systems, the reverse holds. Furthermore, our results suggest that the behavioral effects of policy interventions such as payment system reform can be different depending on the status quo from which a reform starts.

Acknowledgements.

We thank Lin Jing for his assistance in conducting both experiments in China and for his very helpful comments and suggestions. We are grateful to Chaoliang Yang for testing the z-tree program as well as for translating all the material, including instructions and the text on the computer screens, into Chinese. We also thank him for his assistance in preparing and conducting the experiments at Shandong University. We are grateful to Jingyi Luo for her assistance in translating material from Chinese into English. We thank Tom Chang for his valuable comments and advice given at ASHECON in Los Angeles, USA, in June 2014. We have also benefited from discussions with Daniel Wiesen, participants at the IHEA-World Congress in Sydney in 2013, ASHECON 2014, and the Second Workshop on Behavioral and Experimental Health Economics in Hamilton, Canada, 2015. Thanks to the authors of Hennig-Schmidt et al. (2011) for providing us with their data and experimental design. We are grateful for financial support from both the Independent Innovation Fund of Shandong University (grant no. 2012JC038) and the National Natural Science Foundation of China (grant no. 71373146) for funding the experiments conducted at Shandong University. Financial support for Geir Godager (Project-No. 231776), Heike Hennig-Schmidt (Project-No. 231776) and Tor Iversen from the Research Council of Norway is gratefully acknowledged.

Literature

Abbink, K. and Rockenbach, B. (2006): Option pricing by students and professional traders: A behavioural investigation. *Managerial and Decision Economics*, 27, 497–510. doi:10.1002/mde.1284

Anderson GF, Chalkidou K. (2008): Spending on medical care: More is better?. *the Journal American Medical Association.*, 299, 2444–2445. doi:10.1001/jama.299.20.2444

Brislin, R.W. (1970): Back-translation for cross-cultural research. *Journal of Cross-cultural Psychology*, 1, 185–216.

Brosig-Koch, J., H. Hennig-Schmidt, N. Kairies and D. Wiesen (2017a): The effects of introducing mixed payment systems for physicians—Experimental evidence, *Health Economics*, 26, 243–262 doi: 10.1002/hec.3292

Brosig-Koch, J., H. Hennig-Schmidt, N. Kairies, Kokott, J. and D. Wiesen (2017b): Physician performance pay: Experimental evidence, paper presented at iHEA World Congress, Boston 2017.

Brosig-Koch, J. Heike Hennig-Schmidt, Nadja Kairies-Schwarz, Daniel Wiesen (2016): Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision, *Journal of Economic Behavior & Organization*, 131, Part B, 17–23, doi: 10.1016/j.jebo.2015.04.011

Buckley, N., K. Cuff, J. Hurley, S. Mestelman, S. Thomas, D. Cameron (2015): Support for public provision of a private good with top-up and opt-out: A controlled laboratory experiment, *Journal of Economic Behavior & Organization* 111, 177–196.

Buckley, N., K. Cuff, J. Hurley, S. Mestelman, S. Thomas, D. Cameron (2016): Should I stay or should I go? Exit options with mixed systems of public and private health care finance, *Journal of Economic Behavior & Organization* 131, Part B, November 2016, 62—77.

Cardenas, J.-C. (2005): Bringing the lab to the field and the field to the lab: More than changing subjects, Working paper, Economics Department, Universidad de Los Andes, Bogota, Colombia.

Camerer, C. F. (2003): Behavioral game theory—Experiments in strategic interaction, Princeton University Press, Princeton, New Jersey.

Choudhry NK, Fletcher RH, Soumerai SB. (2005): Systematic review: The relationship between clinical experience and quality of health care. *Annals of Internal Medicine*, 142, 260–273. doi: 10.7326/0003-4819-142-4-200502150-00008

Chen, Xiao-Yang (2007): Defensive medicine or economically motivated corruption? A Confucian reflection on physician care in China today, *Journal of Medicine and Philosophy*, 32, 635–648

Cooper, D. J., J. H. Kagel, W. Lo, and Qi. L. Gu. (1999): Gaming against mananagers in incentive systems: Experimental results with Chinese students and Chinese managers."*American Economic Review*, 89, 781–804.

Costa A., A. Foucart, I. Arnon, M. Aparici, J. Apesteguia (2014a): “Piensa” twice: On the foreign language effect in decision making, *Cognition* 130, 236–254. doi: 10.1016/j.cognition.2013.11.010

Costa A, A. Foucart, S. Hayakawa, M. Aparici, J. Apesteguia, J. Heafner, B. Keysar (2014b): Your morals depend on language. *PLoS ONE*, 9, 4. doi:10.1371/journal. pone.0094842

Cox, J.C., E.P. Green, H. Hennig-Schmidt (2016a): Editorial: Experimental and behavioral economics of healthcare, *Journal of Economic Behavior & Organization* 131, Part B, A1–A4. doi: 10.1016/j.jebo.2016.10.011

Cox, J.C., V. Sadiraj, K. E. Schnier, and J. F. Sweeney (2016b): Incentivizing cost-effective reductions in hospital readmission rates, *Journal of Economic Behavior & Organization*, 131, 24–35. doi: 10.1016/j.jebo.2015.03.014

CPC Central Committee and the State Council (2009): Opinions of the CPC Central Committee and the State Council on Deepening the Health Care System Reform. http://www.china.org.cn/government/scio-press-conferences/2009-04/09/content_17575378.htm.

Devlina, R., Sarma, S., (2008): Do physician remuneration schemes matter? The case of Canadian family physicians. *Journal of Health Economics* 27, 1168–1181.

Eckel, C. and P. Grossman (1996): Altruism in anonymous dictator games, Games and economic behavior, 16, 181–191.

Eggleston, K. (2012): Health care for 1.3 billion: China's remarkable work in progress, Milken Institute Review, Second Quarter 2012: 16–27.

Eggleston, K. (2014): Testimony before the U.S.-China Economic and Security Review Commission Hearing on China's Healthcare Sector, Drug Safety, and the U.S.-China Trade in Medical Products.

Eggleston, K., Ling, L., Qingyue, M., Lindelow, M., & Wagstaff, A. (2008): Health service delivery in China: A literature review. *Health Economics*, 17, 2, 49–65.

Eggleston, K. and W. Yip (2004): Hospital competition under regulated prices: Application to urban health sector reforms in China, International Journal of Health Care Finance and Economics, 4, 343–368.

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

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

Epstein Andrew J, Srinivas Sindhu K, Nicholson Sean, Herrin Jeph, Asch David A (2013): Association between physicians' experience after training and maternal obstetrical outcomes: Cohort study, *BMJ* 346 :f1596.

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

Falk, A., and E. Fehr (2003): Why labor market experiments? *Labor Economics* 10, 399-406.

Falk, A., E. Fehr and C. Zehnder (2006): Fairness perceptions and reservation wages: The behavioral effects of minimum wage laws, *The Quarterly Journal of Economics*, 121, 1347–1381.

Fehr, E., and S. Gächter (2000): Cooperation and punishment in public goods experiments, *American Economic Review*, 90, 341–369.

Fehr, E., and S. Gächter (2002): Altruistic punishment in humans, *Nature*, 415, 137–140.

Fink, C. and N. Kairies-Schwarz (2017):Performance pay in hospitals: an experiment on bonuses and fines. Paper presented at the 2017 Annual meeting of Gesellschaft für experimentelle Wirtschaftsforschung, Kassel Germany.

Fischbacher, U. (2007): Z-tree: Zurich toolbox for readymade economic experiments, *Experimental Economics*, 10, 171–178.

Fischbacher, U. and S. Gächter (2010): Social preferences, beliefs, and the dynamics of free riding in public good experiments, *American Economic Review*, 100, 541–556.

Frechette, G. R. (2015): Laboratory experiments: Professionals versus students, in : Frechette, G. R., & Schotter, A.. Handbook of experimental economic methodology. Oxford, Oxford University Press, 360–390.

Füllbrunn, S. and T. Neugebauer (2013): Limited liability, moral hazard, and risk taking: A safety net game experiment, *Economic Inquiry*, 51, 1389–1403.

Gao, Chen, Fei Xu, Gordon G. Liu (2014): Payment reform and changes in health care in China, *Social Science & Medicine* 111 (2014) 10–16.

Gaynor, M., Gertler, P., (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(6), pp. 1105-1116

Godager, G., H. Hennig-Schmidt, and T. Iversen (2016): Does performance disclosure influence physicians' medical decisions? An experimental study, *Journal of Economic Behavior & Organization*, 131, Part B, pp. 36-46 (dx.doi.org//10.1016/j.jebo.2015.10.005).

Gosden, T., Forland, F., Kristiansen, I., Sutton, M., Leese, B., Guiffrida, A., Sergison, M., Pedersen, L. (2001): Impact of payment method on behavior of primary care physicians: A systematic review. Journal of Health Services Research and Policy 6, 44–54.

Grytten, J., Holst, D., Skau, I., (2009): Incentives and remuneration systems in dental services. *International Journal of Health Care Finance and Economics* 9, 259–278.

Harrison, G. W. and J. A. List (2004): Field Experiments, *Journal of Economic Literature* 42, 1009–1055.

Harrison, G., E. Johnson, M. McInnes and E. E. Rutström (2005): Risk aversion and incentive effects: Comment, *American Economic Review*, 95, 897–904

Heilmann, S. (2008): Policy experimentation in China's economic rise, Studies in Comparative International Development, 43, 1–26. doi: 10.1007/s12116-007-9014-4

Heinemann, F. and M. G. Kocher (2013): Tax compliance under tax regime changes, *International Tax and Public Finance*, 20, 225–246

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 (dx.doi.org/10.1016/j.socscimed.2014.03.001).

Hennig-Schmidt, H., R. Selten, and D. Wiesen (2011): How payment systems affect physicians' provision behavior - An experimental investigation, *Journal of Health Economics*, 30, 367–346.

Herrmann, B., Ch. Thöni, S. Gächter (2008): Anti-social punishment across societies, Science 319, 1362–1367.

Hilgers, D. and M. Wibral (2014): How malleable are choice brackets?—The case of myopic loss aversion. Discussion paper.

Hsiao, William C.L. (1995): The Chinese health care system: Lessons for other nations, *Social Science & Medicine*, 41, 1047–1055.

Hu, S., S. Mestelman and W. Scarth (2015): Variable notions of fairness: Experimental evidence from an efficiency-wage environment, McMaster University Department of Economics Working Paper 2015–05

Iversen, T., Lurås, H. (2000): The effect of capitation on GPs' referral decision. *Health Economics* 9, 199–210.

Iversen, T. (2004): The effects of a patient shortage on general practitioners' future income and list of patients*Journal of Health Economics* 23 (4), 673-694

Keser, C., E. Peterle and C. Schnitzler (2014): Money talks—paying physicians for performance, CIRANO Scientific Series 2014s-41.

Keser, C., C. Montmarquette, M. Schmidt and C. Schnitzler (2013): Custom-made healthcare—An experimental investigation, CIRANO Scientific Series 2013s–15.

Levene, H. (1960): Robust tests for equality of variances. In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, ed. I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, and H. B. Mann, 278–292. Menlo Park, CA: Stanford University Press.

McAlister, F. A., Youngson, E., Bakal, J. A., Holroyd-Leduc, J., & Kassam, N. (2015): Physician experience and outcomes among patients admitted to general internal medicine teaching wards. *Canadian Medical Association Journal*, 187(14), 1041–1048. doi: 10.1503/cmaj.150316

MoHRSS, Ministry of Human Resources and Social Security (2011): Opinions of MOHRSS on advancing the health insurance payment method reform. Document No. 63 (in Chinese).

MoHRSS, Ministry of Human Resources and Social Security (2013): 2012 China Labor Statistics Yearbook.

Neugebauer, T. and J. Perote (2008): Bidding 'as if' risk neutral in experimental first price auctions without information feedback, *Experimental Economics*, 11: 190–202.

Newhouse, J. (1996): Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection, *Journal of Economic Literature*, 34,(3), 1236-1263.

Nikiforakis, N. (2008): Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics*, 92, 91–112.

Owens, M., and J. Kagel (2010): Minimum wage restrictions and employee effort in incomplete labor markets: An experimental investigation, *Journal of Economic Behavior & Organization*, 73, 3, 217–226.

Pham HH, Schrag D, Hargraves JL, Bach PB. (2005): Delivery of preventive services to older adults by primary care physicians. *the Journal American Medical Association*. 294, 4, 473–481.

Ramesh, M. and Xun Wu (2009): Health policy reform in China: Lessons from Asia, *Social Science & Medicine* 68, 2256–2262

Reid RO, Friedberg MW, Adams JL, McGlynn EA, Mehrotra A. (2010): Associations between physician characteristics and quality of care. Arch Intern Med. 170, 16, 1442–1449. doi:10.1001/archinternmed.2010.307

Scott A. (2000): Economics of general practice. In: Newhouse J.P., Culyer A,J. (eds.): Handbook of Health Economics. North Holland, Amsterdam, 2000

Scott, A., Witt, J., Humphreys, J., Joyce, C., Kalb, G., Jeon, S.H. and McGrail, M. (2013): Getting doctors into the bush: General practitioners' preferences for rural location, Social Science & Medicine 96, 33–44.

Sørensen, R., Grytten, J., (2003): Service production and contract choice in primary physician services. *Health Policy* 66, 73–93.

Stoddard, B. (2015): Probabilistic production of a public good, *Economics Bulletin*, 35, 1, 37¬–52.

Streja DA, Rabkin SW (1999): Factors associated with implementation of preventive care measures in patients with diabetes mellitus, *Archives of Internal Medicine*, 159, 3, 294–302

The State Council (2009): Essential Implementation Schemes on Health Care System Reform (2009 - 2011) (in Chinese). http://www.gov.cn/zwgk/2009-04/07/content_1279256.htm.

Wang, Hong, Licheng Zhang, Winnie Yip and William Hsiao (2011): An experiment in payment reform for doctors in rural China reduced some unnecessary care but did not lower total osts, *Health Affairs*, 30, 2427–2436.

Weinberger SE, Duffy FD, Cassel CK. (2005): "Practice makes perfect" . . . Or does it?. Ann Intern Med. 142, 302–303. doi: 10.7326/0003-4819-142-4-200502150-00014

Yip, Winnie and William C. Hsiao, (2008): The Chinese health system at a crossroads, *Health Affairs*, 27, 460–468.

Yip, Winnie and W. C. Hsiao (2009): China's health care reform: A tentative assessment, *China Economic Review*, 20, 613–619.

Yip, Winnie, William Hsiao, Qingyue Meng, Wen Chen, Xiaoming Sun (2010): Realignment of incentives for health-care providers in China, *Lancet* 375: 1120–30.

Yip, Winnie, et al. (2014): Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural China. Health Affairs, 33.3: 502-510.

Appendix A

A. Experimental parameters.

Table A1: Experimental parameters

	Payment	Var	0	1	2	3	4	5	6	7	8	9	10
I	FFS	R_{jA}(q)	0.00	1.70	3.40	5.10	5.80	10.50	11.00	12.10	13.50	14.90	16.60
		$R_{jB}(q)$	0.00	1.00	2.40	3.50	8.00	8.40	9.40	16.00	18.00	20.00	22.50
		$R_{jC}(q)$	0.00	1.80	3.60	5.40	7.20	9.00	10.80	12.60	14.40	16.20	18.30
		$R_{jD}(q)$	0.00	2.00	4.00	6.00	8.00	8.00	15.00	16.90	18.90	21.30	23.60
		$R_{jE}(q)$	0.00	1.00	2.00	6.00	6.70	7.60	11.00	12.30	18.00	20.50	23.00
	CAP	$R(q)$	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00
II	FFS,CAP	c(q)	0.00	0.10	0.40	0.90	1.60	2.50	3.60	4.90	6.40	8.10	10.00
III	FFS	$\pi_{jA}(q)$	0.00	1.60	3.00	4.20	4.20	8.00	7.40	7.20	7.10	6.80	6.60
		$\pi_{jB(q)}$	0.00	0.90	2.00	2.60	6.40	5.90	5.80	11.10	11.60	11.90	12.50
		$\pi_{jC(q)}$	0.00	1.70	3.20	4.50	5.60	6.50	7.20	7.70	8.00	8.10	8.30
		$\pi_{jD(q)}$	0.00	1.90	3.60	5.10	6.40	5.50	11.40	12.00	12.50	13.20	13.60
		$\pi_{jE(q)}$	0.00	0.90	1.60	5.10	5.10	5.10	7.40	7.40	11.60	12.40	13.00
	CAP	$\pi(q)$	12.00	11.90	11.60	11.10	10.40	9.50	8.40	7.10	5.60	3.90	2.00
IV	FFS,CAP	B1k(q)	0.00	0.75	1.50	2.00	7.00	10.00	9.50	9.00	8.50	8.00	7.50
		$B_{2k}(q)$	0.00	1.00	1.50	10.00	9.50	9.00	8.50	8.00	7.50	7.00	6.50
		$B_{3k}(q)$	0.00	0.75	2.20	4.05	6.00	7.75	9.00	9.45	8.80	6.75	3.00

Note: This table shows all experimental parameters. $R_{jk}(q)$ denotes physicians' payment for patient type j and illness k. Under FFS, $R_{jk}(q)$ varies with illnesses k and increases in q, whereas under CAP, $R_{jk}(q)$ remains constant. The costs for providing medical services $c_{jk}(q)$ increase in q and are the same under all experimental conditions. The physicians' profit $\pi_{jk}(q)$ is equal to $R_{jk}(q) - c_{jk}(q)$. $B_{jk}(q)$ denotes the patient benefit for the three patient types $j = 1, 2, 3$ held constant across conditions.

Appendix B

Table B1: P-values from Mann-Whitney U tests on differences between Chinese medical students and doctors in providing patient benefit, $B(q)$ and quantity, q

H_0: q ($B(q)$) for patient j do not differ between Chinese medical students and doctors under payment system P.

	Patient benefit, $B(q)$		Service provision (q)	
Patient	FFS	CAP	FFS	CAP
1A	0.0006***	0.0146*	0.5781	0.9020
1B	0.1901	0.0018**	0.1282	0.3572
1C	0.5429	0.0028**	0.9716	0.4884
1D	0.2557	0.0022**	0.4986	0.4630
1E	0.3944	0.0026**	0.9209	0.4921
2A	0.3162	0.0001***	0.9073	0.0047**
2B	0.8363	0.0006***	0.7379	0.0056**
2C	0.2321	0.0056**	0.3470	0.0224*
2D	0.3149	0.0001***	0.5549	0.0009***
2E	0.0712	0.0004***	0.2765	0.0173*
3A	0.0024**	0.2647	0.1378	0.3730
3B	0.0152*	0.1696	0.0513*	0.3713
3C	0.0019**	0.0522	0.0168*	0.2028
3D	0.0011**	0.0030**	0.0024**	0.0552*
3E	0.2327	0.1447	0.0002***	0.1106

Note: This table shows p-values from Mann-Whitney U tests (two-sided),pooled over both parts of the experiment (N=277). With a Bonferroni correction, adusting significance level with a factor of 1/15 (i.e. p=0.0033), we no longer reject the hypothesis that a difference exists in patient benefit for patient 3B under FFS, as well as patients 1A and 2C under CAP. For quantity decisions significance stays for patient 3D and 3E under FFS and for patient 2D under CAP.

Table B2: P-values from Mann-Whitney U tests on Impact of payment system sequence on providing patient benefit, $B(q)$ and quantity, q.

H_0: q ($B(q)$) for patient j under payment system P is not affected by whether P was implemented first or second.

	Patient benefit, B(q)		Service provision (q)	
Patient	FFS	CAP	FFS	CAP
1A	0.0542	0.0003***	0.3488	0.0038**
1B	0.1655	0.0003***	0.7282	0.0679*
1C	0.1470	<0.0001***	0.2979	0.6513
1D	0.2749	0.1035	0.3693	0.7018
1E	0.3497	0.0004***	0.5875	0.6282
2A	0.7358	<0.0001***	0.9118	0.0985*
2B	0.3102	<0.0001***	0.3079	<0.0001***
2C	0.7395	0.0001***	0.9344	0.0030**
2D	0.0294*	<0.0001***	0.0164*	0.0113*
2E	0.4146	0.0009***	0.2067	0.0189*
3A	0.2101	0.2664	0.2348	0.7891
3B	0.2420	0.2494	0.9379	0.7503
3C	0.8454	0.3660	0.7435	0.9744
3D	0.3128	0.2314	0.3128	0.9814
3E	0.3447	0.3341	0.1880	0.7582

Note: This table shows p-values from Mann-Whitney U tests (two-sided),pooled over both parts of the experiment (N=277). With a Bonferroni correction, adusting significance level with a factor of 1/15 (i.e. p=0.0033), patient benefit for patient 2D under FFS becomes insignificant. For quantity decisions significance stays for patients 2B and 2C under CAP.

Table B3: P-values from Mann-Whitney U tests on differences between between German and Chinese medical students in providing patient benefit, $B(q)$ and quantity, q.

H_0: q ($B(q)$) for patient j does not differ between German and Chinese medical students under payment system P.

	Patient benefit, B(q)		Service provision (q)	
Patient	FFS	CAP	FFS	CAP
1A	0.0723	0.6697	0.8900	0.1297
1B	0.7292	0.4569	0.1817	0.5609
1C	0.4745	0.9134	0.5201	0.2318
1D	0.8477	0.4318	0.8981	0.4528
1E	0.1164	0.9242	0.0592	0.5289
2A	0.6086	0.9239	0.4817	0.6205
2B	0.5968	0.3330	0.4899	0.2389
2C	0.2896	0.0748	0.1691	0.0764
2D	0.2875	0.2835	0.2213	0.0414*
2E	0.2622	0.2264	0.1990	0.0891
3A	0.8754	0.5438	0.5748	0.9969
3B	0.7403	0.9418	0.4671	0.7383
3C	0.0968	0.7697	0.0554	0.9019
3D	0.2142	0.9598	0.2252	0.7067
3E	0.4281	0.2509	0.3043	0.0799

Note: With a Bonferroni correction, adusting significance level with a factor of 1/15 (i.e. p=0.0033), service provision for patient 2D under CAP becomes insignificant.

Table B4: P-values from chi^2 tests on differences between German and Chinese medical students in probability of providing $q = q^*$ H_0: The proportion of patient benefit-maximizing subjects for patient j does not differ between German and Chinese medical students under payment system P.

Patient	FFS	CAP
1A	0.070*	1.000
1B	0.958	0.434
1C	0.683	0.620
1D	0.816	0.483
1E	0.472	1.000
2A	0.958	1.000
2B	0.600	0.284
2C	0.801	0.060*
2D	0.766	0.216
2E	0.785	0.182
3A	0.927	0.693
3B	0.801	0.835
3C	0.133	0.609
3D	0.194	1.000
3E	0.475	0.480

Note: With a Bonferroni correction, adusting significance level with a factor of 1/15 (i.e. p=0.0033), service provision for patient 1A under FFS and 2C under CAP become insignificant.

Table B5: Regressions results comparing behavioral results from Chinese doctors and medical students with German medical students.

	Ordinal regression model dependent variable: q			Ordinal regression model dependent variable: B(q)			Logistic regression model for the probability of choosing $q = q^*$		
Regressors	Odds Ratio	95% C.I.		Odds Ratio	95% C.I.		Odds Ratio	95% C.I.	
FFS	11.66	8.25	16.48	0.73*	0.52	1.02	0.44***	0.31	0.64
Doctor	0.98	0.65	1.48	0.62**	0.43	0.89	0.51***	0.34	0.75
Germany	1.17	0.72	1.90	0.98	0.58	1.67	1.12	0.61	2.07
1B	3.26***	2.53	4.20	0.16***	0.10	0.26	0.15***	0.10	0.24
1C	2.77***	2.18	3.52	0.25***	0.16	0.38	0.16***	0.11	0.24
1D	3.83***	3.01	4.87	0.24***	0.15	0.38	0.13***	0.09	0.21
1E	4.77***	3.52	6.48	0.20***	0.12	0.33	0.15***	0.10	0.23
2A	0.29***	0.22	0.38	0.25***	0.15	0.42	0.15***	0.10	0.24
2B	0.80	0.56	1.16	0.24***	0.13	0.42	0.15***	0.09	0.24
2C	0.52***	0.36	0.75	0.31***	0.17	0.54	0.26***	0.16	0.40
2D	0.86	0.58	1.27	0.20***	0.11	0.37	0.20***	0.13	0.32
2E	0.68	0.42	1.10	0.33***	0.18	0.63	0.41***	0.26	0.63
3A	5.76***	4.36	7.62	0.05***	0.03	0.08	0.10***	0.07	0.15
3B	8.22***	6.21	10.87	0.06***	0.04	0.10	0.23***	0.16	0.33
3C	9.62***	7.29	12.70	0.06***	0.04	0.10	0.18***	0.12	0.25
3D	10.19***	7.69	13.49	0.06***	0.04	0.10	0.22***	0.15	0.32
3E	12.58***	9.14	17.32	0.05***	0.03	0.08	0.07***	0.05	0.11
Log pseudolikelihood	-7862.3			-10087			-2689.1		
#obs	4785								
# individuals	319								
#obs per individual	15								

Estimates with ***(**)[*] indicate that the odds ratio is significantly different from 1 at the 0.1 % (1 %) [5 %] level in a two-tailed test.

Appendix C

C1: Instructions of the experiment

[Numbers/text in brackets refer to the conditions where doctors participate.]

{Sentences/decision screens in braces are inserted into the instructions either in condition FFS or in condition CAP.}

Instructions Part 1
General Information
In the following experiment, you will make a couple of decisions. Following the instructions and depending on your decisions, you can earn money. It is therefore very important that you read the instructions carefully.

You take your decisions anonymously on your computer screen. During the experiment, you are not allowed to talk to any other participant. Whenever you have a question, please raise your hand. The experimenter will answer your question in private in your cubicle. If you disregard these rules, you can be excluded from the experiment without receiving any payment. All amounts of money in the experiment are stated in Token. At the end of the experiment, your earnings will be converted into RMB at an exchange rate of 10 Token = 1 [6] RMB and paid to you in cash.

The experiment consists of two parts. We we will inform you now on the decision situation in Part 1. We will provide you with the instructions for Part 2 as soon as Part 1 has ended. Please note that your decisions in Part 1 have no influence on your decisions in Part 2 and vice versa.

Your decisions in Part 1 of the experiment
During the experiment, you are in the role of a physician. You have to make 15 decisions regarding the treatment of patients. All participants of this experiment take their decisions in the role of physicians. You decide on the quantity of medical services you want to provide for given clinical symptoms of a patient.

You decide on your computer screen where five different kinds of clinical symptoms – A, B, C, D, and E – of three different patient types – 1, 2, and 3 – will be shown one after another. For each patient you can provide 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 medical services.

Your remuneration is as follows:

Condition CAP: For each patient you receive a lump-sum payment that is independent of the quantity of medical services.

Condition FFS: A different payment is assigned to each quantity of medical services. The payment increases in the quantity of medical services.

While deciding on the quantity of medical services, in addition to your payment you determine the costs you incur when providing these services. Costs increase with increasing quantity provided. Your profit in Token is calculated by subtracting your costs from your payment.

A certain benefit for the patient is assigned to each quantity of medical services, the patient benefit that the patient gains from your provision of services (treatment). Therefore, your decision on the quantity of medical services not only determines your own profit, but also the patient benefit. An example for a decision situation is given on the following screen.

{Decision screen for patient 1C under FFS}

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	0.00	0.00	0.00	0.00
Service C1	1	1.80	0.10	1.70	0.75
Service C1, Service C2	2	3.60	0.40	3.20	1.50
Service C1, Service C2, Service C3	3	5.40	0.90	4.50	2.00
Service C1, Service C2, Service C3, Service C4	4	7.20	1.60	5.60	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	9.00	2.50	6.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	10.80	3.60	7.20	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.60	4.90	7.70	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	14.40	6.40	8.00	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	16.20	8.10	8.10	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	18.30	10.0	8.30	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

{Decision screen for patient 1C under CAP}

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	12.00	0.00	12.00	0.00
Service C1	1	12.00	0.10	11.90	0.75
Service C1, Service C2	2	12.00	0.40	11.60	1.50
Service C1, Service C2, Service C3	3	12.00	0.90	11.10	2.00
Service C1, Service C2, Service C3, Service C4	4	12.00	1.60	10.40	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	12.00	2.50	9.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	12.00	3.60	8.40	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.00	4.90	7.10	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	12.00	6.40	5.60	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	12.00	8.10	3.90	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	12.00	10.0	2.00	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

You decide on the quantity of medical services on your computer screen by typing an integer between 0 and 10 into the box labeled "Your Decision".

After all participants have taken their decisions for the respective patient you will proceed to the next patient. There are no real, but abstract patients participating in this experiment. Yet, the patient benefit, which an abstract patient receives by your providing medical services, will be beneficial for a real patient. The total amount of patient benefit determined by your 15 decisions will be provided to a patient with cancer treated in Shandong Qilu Hospital [Shandong Provincial Cancer Hospital]. The money will be directly transferred to the patient's account in the hospital, to help him/her with part of the treatment fee.

Each time you make a decision on the quantity of medical services you will be informed on your profit and the patient benefit. After you have made your 15 decisions in Part 1 of the experiment you will get to know your total profit and the corresponding total patient benefit.

Earnings in Part 1 of the experiment

After you have made your decisions in Part 1 of the experiment, your overall earnings will be calculated by summing up your profits from providing medical services to the 15 patients. This amount will be converted from Token into RMB. Your earnings of Part 1 of the experiment together with the earnings of Part 2 will be paid to you in cash at the end of the experiment (rounded to 1 Yuan).

The patient benefit gained by all 15 patients will be converted into RMB at the end of the experiment, too, and will be transferred to the real patient's account. To this end the experimenter and a monitor will go together to Shandong Qilu Hospital [Shandong Provincial Cancer Hospital]. After the transfer, the signed receipt will be scanned into electronic form and will be sent to all the participants via e-mail in order to ensure the authenticity of the above process. Personal information will be blinded black to respect the patient's privacy.

After the end of Part 2 of the experiment, one participant is randomly assigned the role of the monitor. The monitor receives a payment of 50 [200] RMB in addition to the payment from the experiment. In the end, the monitor signs a form to verify that the procedure described above was actually carried out. This form will be sent to all participants together with the receipt via e-mail.

Next, please answer some questions familiarizing you with the decision situation. After your 15 decisions, please answer some further questions on your screen.

Instructions Part 2

The experiment will now be repeated including one change. Like in Part 1 you will make 15 decisions. After these 15 decisions the experiment will end.

The General Information from Part 1 also applies for Part 2 of the experiment.

Your decisions in Part 2 of the experiment
Also in Part 2 of the experiment, you are in the role of a physician and you have to make 15 decisions regarding the treatment of patients. All participants take their decisions in the role of physicians. You decide on the quantity of medical services you want to provide for given clinical symptoms of a patient.

Like in Part 1 you decide on your computer screen where five different kinds of clinical symptoms A, B, C, D, and E of three different patient types (1, 2, and 3) will be shown one after another. For each patient you can provide 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 medical services.

Your remuneration is as follows:
{Condition CAP: For each patient you receive a lump-sum payment that is independent of the quantity of medical services.}
{Condition FFS: A different payment is assigned to each quantity of medical services. The payment increases in the quantity of medical services.}

As in Part 1, while deciding on the quantity of medical services, in addition to your payment you determine the costs you incur when providing these services. Costs increase with increasing quantity provided. Your profit in Token is calculated by subtracting your costs from your payment.

A certain benefit for the patient is assigned to each quantity of medical services, the patient benefit that the patient gains from your provision of services (treatment). Therefore, your decision on the quantity of medical services not only determines your own profit, but also the patient benefit. An example for a decision situation is given on the following screen.

{Decision screen for patient 1C under FFS}

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	0.00	0.00	0.00	0.00
Service C1	1	1.80	0.10	1.70	0.75
Service C1, Service C2	2	3.60	0.40	3.20	1.50
Service C1, Service C2, Service C3	3	5.40	0.90	4.50	2.00
Service C1, Service C2, Service C3, Service C4	4	7.20	1.60	5.60	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	9.00	2.50	6.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	10.80	3.60	7.20	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.60	4.90	7.70	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	14.40	6.40	8.00	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	16.20	8.10	8.10	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	18.30	10.0	8.30	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

{Decision screen for patient 1C under CAP}

Patient type 1/Illness C

Medical services	Quantity	Your Remuneration (in Taler)	Your Cost (in Taler)	Your Profit (in Taler)	Patient benefit (in Taler)
none	0	12.00	0.00	12.00	0.00
Service C1	1	12.00	0.10	11.90	0.75
Service C1, Service C2	2	12.00	0.40	11.60	1.50
Service C1, Service C2, Service C3	3	12.00	0.90	11.10	2.00
Service C1, Service C2, Service C3, Service C4	4	12.00	1.60	10.40	7.00
Service C1, Service C2, Service C3, Service C4, Service C5	5	12.00	2.50	9.50	10.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6	6	12.00	3.60	8.40	9.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7	7	12.00	4.90	7.10	9.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8	8	12.00	6.40	5.60	8.50
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9	9	12.00	8.10	3.90	8.00
Service C1, Service C2, Service C3, Service C4, Service C5 Service C6, Service C7, Service C8, Service C9, Service C10	10	12.00	10.0	2.00	7.50

Your Decision

Please indicate the quantity of medical services you want to provide

OK

You decide on the quantity of medical services on your computer screen by typing an integer between 0 and 10 into the box labeled ”Your Decision”.

After all participants have taken their decisions for the respective patient you will proceed to the next patient.

Also in this part of the experiment there are no real, but abstract patients participating in this experiment. Yet, the patient benefit, which an abstract patient receives by your providing medical services, will be beneficial for a real patient. Also in the second part of the experiment the total amount of patient benefit determined by your 15 decisions will be provided to a patient with cancer treated in Shandong Qilu Hospital [Shandong Provincial Cancer Hospital]. The money will be directly transferred to the patient's account in the hospital, to help him/her with part of the treatment fee.

Each time you made a decision on the quantity of medical services you will be informed on your profit and the patient benefit. After you have made your 15 decisions in Part 2 of the experiment you will get to know your total profit and the corresponding total patient benefit.

Earnings in Part 2 of the experiment
After you have made your decisions in Part 2 of the experiment, your overall earnings will be calculated by summing up your profits from providing medical services to the 15 patients. This amount will be converted from Token into RMB at the end of the experiment and will be paid to you in cash together with the earnings of Part 1 of the experiment (rounded to 1 Yuan).

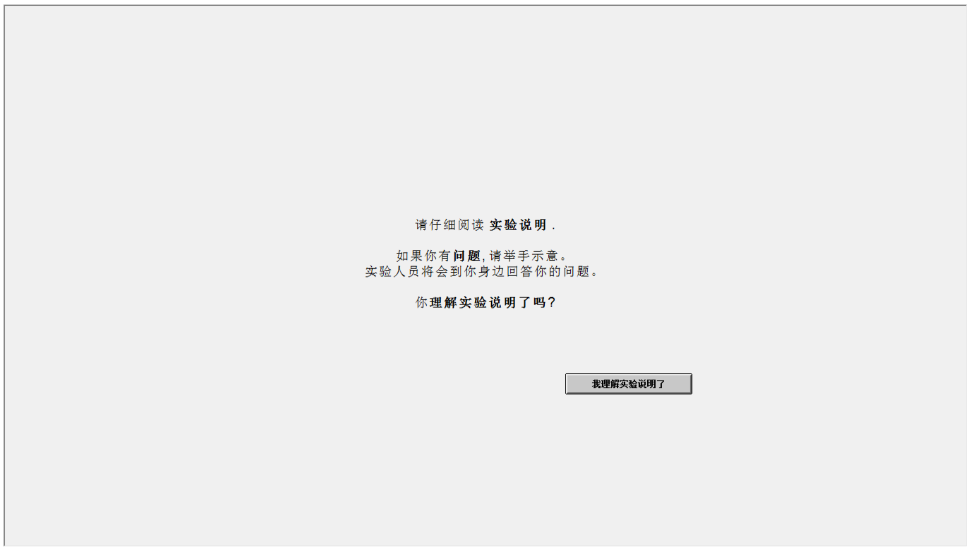
The patient benefit gained by all 15 patients will be converted into RMB at the end of the experiment, too, and will be transferred to the real patient's account. To this end the experimenter and a monitor will go together to Shandong Qilu Hospital [Shandong Provincial Cancer Hospital]. After the transfer, the signed receipt will be scanned into electronic form and will be sent to all the participants via e-mail in order to ensure the authenticity of the above process. Personal information will be blinded black to respect the patient's privacy. Information about the procedure has been given in Part 1 of the experiment.

Next, please answer some questions in this part of the experiment that will familiarize you with the present decision situation. After your 15 decisions, please answer some further questions on your screen.

C2: Test questions prior to the experiment

The following example applies to FFS condition. For CAP condition, screens 3 to 5 are similar to Figure D1b in Appendix D.

Screen 1



Please read the instructions carefully. If you have a question, please raise your hand. The experimenter will come to you and answer your question. Have you understood the instructions?

Screen 2

To familiarize you with the decision situation we first ask you to answer 3 questions. We will inform you when the actual experiment starts.

Screen 3 [4, 5]

患者类型 1 / 临床症状 F

医疗服务	数量	你的诊疗费 (以代币计算)	你的成本 (以代币计算)	你的净收益 (以代币计算)	患者效益 (以代币计算)
不提供	0	0.00	0.00	0.00	0.00
服务 F1	1	0.90	0.10	0.80	0.75
服务 F1 / 服务 F2	2	1.60	0.40	1.20	1.50
服务 F1 / 服务 F2 / 服务 F3	3	5.10	0.90	4.20	2.00
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4	4	5.10	1.60	3.50	7.00
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5	5	5.10	2.50	2.60	10.00
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5 / 服务 F6	6	7.40	3.60	3.80	9.50
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5 / 服务 F6 / 服务 F7	7	7.40	4.90	2.50	9.00
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5 / 服务 F6 / 服务 F7 / 服务 F8	8	11.60	6.40	5.20	8.50
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5 / 服务 F6 / 服务 F7 / 服务 F8 / 服务 F9	9	12.40	8.10	4.30	8.00
服务 F1 / 服务 F2 / 服务 F3 / 服务 F4 / 服务 F5 / 服务 F6 / 服务 F7 / 服务 F8 / 服务 F9 / 服务 F10	10	13.00	10.00	3.00	7.50

假设一位医生准备为上述患者提供数量为0 项的医疗服务。

1 a) 诊疗费是多少?

1 b) 成本是多少?

1 c) 净收益是多少?

1 d) 患者效益是多少?

OK

Assume a physician wants to provide the quantity of 0 [10, 4] medical services for the patient above.
1 [2, 3] a) What is the remuneration?
1 [2, 3] b) What are the costs?
1 [2, 3] c) What is the profit?
1 [2, 3] d) What is the patient benefit?

Screen 6 The test questions are now completed. When you click on the button the experiment will start!

C3: Questionnaires after Part 1 and Part 2 of the experiment.

[Information in brackets was requested from doctors only.]

{{Numbers/text in double braces refer to Part 2 of the experiment.}}

Please confirm your terminal number on your questionnaire. After you have made all decisions in Part 1 {{2}} of the experiment we would like to ask you to answer the following questions as good as possible. These answers are extremely important for our studies. Thank you for your cooperation.

Please put yourself back into the decision situation of Part 1 {{2}} of the experiment.

- What factors did influence your decision? Why did you decide in this way?
- How did the profit influence your decision?
- How did the patient benefit influence your decision?
- {{Major (faculty / main subject(s)):}}
- {{What is the number of your semester?}}
- {{Your gender: female/male}}
- {{Your nationality: (students only)}}
- {{[Your age]:}}
- {{[How many years of professional experience do you have?]}}
- {{[Your specification
- General Practitioner
- Traditional Chinese Medicine
- Public Health
- Other]}}

D. Chinese decision screens.

Figure D1a. Illustration of the decision screen for patient 1C under CAP

患者类型 1/临床症状 C

医疗服务	数量	你的诊疗费 (代币)	你的成本 (代币)	净收益 (代币)	患者效益 (代币)
不提供	0	0.00	0.00	0.00	0.00
服务 C1	1	1.80	0.10	1.70	0.75
服务 C1, 服务 C2	2	3.60	0.40	3.20	1.50
服务 C1, 服务 C2, 服务 C3	3	5.40	0.90	4.50	2.00
服务 C1, 服务 C2, 服务 C3, 服务 C4	4	7.20	1.60	5.60	7.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5	5	9.00	2.50	6.50	10.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6	6	10.80	3.60	7.20	9.50
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7	7	12.60	4.90	7.70	9.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8	8	14.40	6.40	8.00	8.50
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8, 服务 C9	9	16.20	8.10	8.10	8.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8, 服务 C9, 服务 C10	10	18.30	10.0	8.30	7.50

你的决策

请填写你要提供的医疗服务的数量

OK

Figure D1b. Illustration of the decision screen for patient 1C under FFS

患者类型 1/临床症状 C

医疗服务	数量	你的诊疗费 (代币)	你的成本 (代币)	净收益 (代币)	患者效益 (代币)
不提供	0	12.00	0.00	12.00	0.00
服务 C1	1	12.00	0.10	11.90	0.75
服务 C1, 服务 C2	2	12.00	0.40	11.60	1.50
服务 C1, 服务 C2, 服务 C3	3	12.00	0.90	11.10	2.00
服务 C1, 服务 C2, 服务 C3, 服务 C4	4	12.00	1.60	10.40	7.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5	5	12.00	2.50	9.50	10.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6	6	12.00	3.60	8.40	9.50
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7	7	12.00	4.90	7.10	9.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8	8	12.00	6.40	5.60	8.50
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8, 服务 C9	9	12.00	8.10	3.90	8.00
服务 C1, 服务 C2, 服务 C3, 服务 C4, 服务 C5 服务 C6, 服务 C7, 服务 C8, 服务 C9, 服务 C10	10	12.00	10.0	2.00	7.50

你的决策

请填写你要提供的医疗服务的数量

OK