Bonn Econon Discussion Papers

Discussion Paper 29/2009

How Payment Systems Affect Physicians´ Provision Behaviour – An Experimental Investigation

by

Heike Henning-Schmidt, Reinhard Selten
and Daniel Wiesen

December 2009

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Kaiserstrasse 1
D-53113 Bonn

Financial support by the
Deutsche Forschungsgemeinschaft (DFG)
through the
Bonn Graduate School of Economics (BGSE)
is gratefully acknowledged.

Deutsche Post World Net is a sponsor of the BGSE.

How Payment Systems Affect Physicians' Provision Behaviour – An Experimental Investigation*

Heike Hennig-Schmidt†, Reinhard Selten† & Daniel Wiesen†‡

July 23, 2009

Abstract

A central concern in health economics is to understand the influence of commonly used physician payment systems. We introduce a controlled laboratory experiment to analyze the influence of fee-for-service (FFS) and capitation (CAP) payments on physicians' behaviour. Medical students decide as experimental physicians on the quantity of medical services. Real patients gain a monetary benefit from their choices. Our main findings are that patients are overserved in FFS and underserved in CAP. Financial incentives are not the only motivation for physicians' quantity decisions, though. The patient benefit is of considerable importance as well. Patients are affected differently by the two payment systems. Those patients in need of a low level of medical services are better off under CAP, whereas patients with a high need of medical services gain more health benefit when physicians are paid by FFS.

Keywords: Physician payment system; laboratory experiment; incentives; fee-for-service; capitation

JEL-Classification: C91, I11

*The present paper is a completely revised version of Hennig-Schmidt, Reinhard Selten & Daniel Wiesen (2007) "Physicians' Provision Behavior under Different Payment Systems – An Experimental Investigation". We are grateful for helpful comments and suggestions by participants of the following conferences and seminars: IZA Workshop on Behavioral and Organizational Economics, Sozialökonomischer Ausschuss des Vereins für Socialpolitik, IMEBE, EEA, IAREP, ESA European Meeting, iHEA World Congress, University of Zurich, HERO at the University of Oslo, cege at the University of Göttingen and Erasmus University Rotterdam. Financial support for Daniel Wiesen by the Konrad-Adenauer-Stiftung e.V. is gratefully acknowledged.

†*B*onn*E*con*L*ab, Laboratory for Experimental Economics, Department of Economics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany

‡Corresponding author: Phone: +49(0)228 739194, Fax: +49(0)228 739193, e-mail: daniel.wiesen@uni-bonn.de.

1 Introduction

A central concern in health economics is to understand the influence of institutions on the behaviour of actors on health care markets. In practice, effects from changing institutions like the payment system during a health care reform are *ex ante* not necessarily known to policy makers and may influence behaviour in an undesired way. Main addressees of reforms are health care providers (physicians) whose behaviour is likely to be influenced by the payment system. Theoretical health-economic literature has highlighted the different incentives of commonly used payment systems like fee-for-service (FFS) or capitation (CAP). Under FFS physicians are paid for each medical procedure or service dispensed to a patient whereas under CAP, physicians receive a fixed payment for each patient irrespective of the quantity of medical services provided. FFS inherits an incentive to 'overserve' patients whereas CAP may lead to underprovision of medical services (Ellis and McGuire 1986, Newhouse 2002).

Field studies show that different payment systems do affect physicians' behaviour. Yet, the results are often not comparable because of country-specific institutional differences.[1] In some studies, more than one component of the payment system are varied simultaneously making causal inferences difficult or even impossible. According to Gosden et al. (2001) the results are too contradictory to draw a definite conclusion on the direction of an effect.

Another empirical method is called for that complements field studies and overcomes (some of) the problems mentioned above. Fuchs (2000) in his article on the future of health economics argues that incorporating methods of experimental economics into health economic research may lead to great benefits for the latter. In a similar vein, Frank (2007) argues in favor of applying behavioural economics methods in health economics.

Our study contributes to the research agendas suggested by Fuchs and Frank. We use a controlled laboratory experiment to improve the understanding of the institutional parameter 'payment system' by implementing the specific features of FFS and CAP. The main focus of our study is on *how* the two payment systems influence a physician's provision of medical services, and we abstract from factors other than the payment system. Our study is one of the very first ones tackling a health economic topic by methods of experimental economics.[2]

In our experiment, experimental physicians – all of them medical students – decide on the quantity of medical services under the two payment systems. Patients gain a benefit from these services, the patient benefit measured in monetary terms. Only abstract patients 'participate' in our experiment. To provide the experimental physicians with an incentive for favourable behaviour towards the patients, however, the money corresponding to the benefits of all abstract patients is transferred to a charity caring for real patients.

Our main finding is that physicians *are* influenced by the payment system. In line with

[1]See for example the studies by Stearns et al. (1992) and Davidson et al. (1992) in the US, Krasnik et al. (1990) in Denmark, Iversen and Lurås (2000) and Grytten and Sørensen (2001) in Norway, Hutchinson et al. (1996), Devlina and Sarma 2008 and Dumont et al., 2008 in Canada.

[2]Other studies are Fan et al. (1998), Ahlert et al. (2008) and Hennig-Schmidt and Wiesen (2009).

theoretical considerations, patients are overserved under FFS and underserved under CAP. Financial incentives are not the only motivation for physicians' quantity decisions, though; the patient benefit is of considerable importance as well. Patients are affected differently by the two payment systems. Those in need of a low level of medical services are better off under CAP, whereas patients needing a high level of medical services gain a higher health benefit when physicians are paid by FFS.

Our paper is organized as follows. Section 2 sketches the theoretical and empirical literature on physician payment and incentives most relevant to our research topic. Section 3 states our research questions. Experimental design and procedure are described in Section 4. Section 5 provides a statistical analysis of subjects' behaviour within and across payment systems. Section 6 concludes.

2 Related literature

In the health economics literature, several authors have highlighted the different incentives in commonly used payment systems like fee-for-service (FFS) or capitation (CAP). Conventional modeling of the physician-patient interaction relies on profit maximization, however (for a summary see McGuire, 2000). Recently, a number of authors depart from modeling physicians as pure profit maximizers by allowing for patient benevolence in the physician's objective function; see for instance Ellis and McGuire (1986), Ma (1994, 2007), Jack (2005) and Choné and Ma (2007).

In their seminal article, Ellis and McGuire (1986) let the physician[3] decide on the quantity of medical services as an agent of the patient and the hospital. The physician's utility derives from two elements – the hospital's profit and the patient's benefit. According to Newhouse (2002), Ellis and McGuire's model is also applicable to a primary care setting rather resembling the setup we are interested in. This implies that the physician is assumed to be concerned about her own profit π and the patient benefit B, both depending on the quantity of medical services q. A major argument for including B into the physician's utility function is the professional code of medical ethics the physician is obliged to (Hippocratic Oath).[4] Ellis and McGuire find that FFS provides an incentive to overserve patients whereas CAP may lead to underprovision of medical services. Moreover, capitation payments can cause underprovision of *necessary* services (Blomqvist, 1991) and may lead to cream-skimming of patients (Newhouse, 1996 and Barros, 2003).

A rich empirical literature has studied various aspects of the relationship between the method of physician remuneration and physician behaviour. Some empirical evidence suggests that physicians do respond to financial incentives. Krasnik et al. (1990) in a before-and-after study, analyse behaviour of general practitioners in Denmark when the system is varied from a (pure) lump-sum payment to CAP supplemented by a FFS component. They find diagnostic and curative services to increase and the number of referrals to secondary care and hospitals to decrease. Concerning referral rates, Iversen and Lurås (2000)

[3]In the following, we denote the physician as female and the patient as male.

[4]See also Arrow (1963) who emphasized the importance of professional ethics; treatment should be determined by objective needs and not be limited by financial considerations.

arrive at a similar result. They analyse referrals from primary to secondary care revealed by Norwegian general practitioners when the payment system was changed from a practice allowance component[5] complemented by a FFS-payment to a CAP-system with a lower FFS-component. The authors find referrals to be larger under CAP (with FFS-component) compared to FFS (with practice-allowance component). The increase in referrals may, however, not only be attributable to CAP but rather to the lower FFS-component.

In a randomized controlled study, Davidson et al. (1992) investigate behaviour of office-based primary care physicians under a FFS system with high and low fees and a CAP system. Patients were children enrolled in the US-*Medicaid* program. Here, the frequency of primary care visits in the high FFS group was higher than in the CAP group. Apparently, CAP physicians constrain the quantity of medical services in order to reduce their costs. The fundholding regulation[6] in CAP may explain the lower referrals to secondary care as the responsibility for children's medical cost seems to outweigh the incentive to minimize cost in CAP.

In a more recent study, Dumont et al. (2008) analyse data on physician services from the Canadian province Quebec before and after a variation from FFS to a mixed system with a base wage, independent of services provided, and a reduced FFS payment. Their results suggest that physicians did react to payment incentives by reducing the volume of (billable) services under the mixed remuneration system. Moreover, these physicians increased the time spent per service and per non-clinical service. The latter are important to insure the quality of health care but are not remunerated under FFS. The results of Dumont et al. suggest a quantity-quality substitution in health care provision.

One of the most important if not the only *controlled field experiment* in health economics is the RAND health insurance study (Newhouse and the Insurance Experiment Group 1993). The main goal of this experiment was to investigate the influence of the insurance system (patients' co-payment vs. free care) on patients' health care service use and their health status. It was found that all types of services analysed in the study fell with cost sharing but the reduced service use had nearly no adverse effect on health for the average person. Health among the sick poor was adversely affected, however. A smaller part of the study was devoted to analyzing the influence of the payment system. To this end, the authors compared the use of services under fee-for-service remuneration with that in a capitated staff model HMO (Health Maintenance Organisation).[7] Cost savings were found to be noticeable, in particular due to lower hospital admission and lower estimated expenditure.

Not all studies support the strong link between physicians' payment systems and their behaviour, however. For example, Hutchinson et al. (1996) do not find differences when comparing hospital utilization rates in Ontario (Canada) under FFS and CAP. For data from Norwegian physicians, Grytten and Sørensen (2001) find that after controlling for characteristics of patients and general practitioners the effects of physicians' payment sys-

[5]A practice allowance is a fixed sum of money Norwegian physicians are paid when contracting with the regional government.

[6]Such a fundholding system has the following characteristics: i) the financial resources for each patient are held in a fund and ii) the general practitioner is usually the decision-maker for allocating the funds.

[7]In a staff model HMO physicians typically work on a salary basis.

tems is rather small.

What can be concluded from the empirical literature cited above? Based on their meta-study, Gosden et al. (2001) acknowledge some empirical evidence that the payment system affects physician behaviour. They stress, however, that field studies face various difficulties like multiple and unobservable influences on physicians' decisions, context and country-specific payment system variations that make the generalization of results difficult. In addition, several field studies suffer from methodological problems when for instance more than one component of the payment system is varied simultaneously. We will return to these issues in the next section.

3 Research questions

Our main research goal is to improve the understanding on how the institutional parameter 'payment system' influences physicians' behaviour. To this end, we make use of experimental economics methods by running a controlled laboratory experiment.

Experimental economics is a valid research technique that can successfully complement field and survey studies. It has a variety of advantages compared to the latter research approaches (see Davis and Holt 1993, Falk and Fehr 2003). Experimental data is created under controlled conditions. It is gathered in experimental sessions in which human subjects supplied with *monetary incentives*[8] make *real* decisions in economically relevant decision situations. Experimental conditions and variables of interest can be varied in a controlled manner. Exogenous *ceteris paribus* variations (e.g. of the payment system) can be easily implemented. Therefore, changes in behaviour can be attributed to these modifications. Different experimenters can repeat the same experiment under comparable conditions to test for the robustness of the results.

Contrary to laboratory data, field data are collected from a natural environment where many factors influence the variable(s) of interest in a way the researcher usually cannot control.[9] These are for instance institutional parameters, physicians' characteristics, uncertainty about the impact of medical services provided as well as patient characteristics like health status or type of insurance. Constant patient populations during a transition of payment systems is important for the validity of results but can most often not be guaranteed. Also, the methodological deficiencies mentioned in the section above should not be neglected (see Gosden et al. 2001). This said, laboratory experimentation apparently is a suitable research method to successfully complement theoretical analyses and other methods of empirical investigation.

Despite the advantages of experimental economics, objections like non-representative student subject pools, low incentives, a small number of observations and the simple environment should be taken seriously. Yet, careful experimentation can avoid many of these problems (see Falk and Fehr 2003).

[8]Participants are paid because they are likely to behave differently when monetary consequences are involved as compared to hypothetical choices (see Camerer and Hogarth 1999 and Hertwig and Ortmann 2001).

[9]See, however, the RAND health insurance experiment (Newhouse and the Insurance Experiment Group 1993)

We are aware that our experiment is very simplistic; in reality, a physician's decision situation is much more complex. Yet, as the goal of the present study is to highlight fundamental consequences of the payment system for physicians' behaviour we think simplicity to be an advantage rather than a deficiency.

The focus of our study is on how the pure payment systems FFS and CAP influence an experimental physician's provision of medical services. We incorporate the two major determinants that according to the theoretical literature referred to in Section 2 influence a physician's behaviour, the own profit and the patient's benefit. We also include patients with different health status, so-called patient types, to account for heterogeneity in the patient population.

Our first research question is concerned with behaviour in FFS. Given our experimental parameters, do experimental physicians tend to behave according to what theory predicts in that they choose a quantity of medical services q^{FFS} larger than the patient's optimal quantity q^* if the profit-maximal quantity $\hat{q}$ exceeds q^*? Taking q^* as the benchmark for the right (best) medical treatment, we conjecture patients to be overserved under FFS.

Second, we are interested in behaviour under CAP. According to predictions from theoretical models we expect patients to be underserved in that physicians choose q^{CAP} lower than q^*.

Third, we are concerned with research questions related to the consequences of *both* payment systems. How does provision behaviour under CAP compare to behaviour under FFS? Based on our previous conjectures, we expect experimental physicians in FFS to choose more medical services than in CAP. Moreover, does the mode of payment have an impact on whether and how experimental physicians besides their own profit take the patient benefit into account? Given the professional code of medical ethics physicians are obliged to, we expect them *not* to behave in a completely self-interested manner.

We also analyse the previous questions with regard to patient types. Does the payment system affect patients with different health status differently as to physicians' treatment? If so, are there differences between FFS and CAP? We expect this to be the case. The RAND health insurance experiment (Newhouse and the Insurance Experiment Group 1993), for instance, showed certain albeit small adverse health consequences concentrated among sick people from the lowest income group.

The last research question concerns the tradeoff between own profit and patient benefit the experimental physicians are faced with. In our experiment, several Pareto-efficient quantity decisions exist for each patient. Here, physicians can neither make the patient better off without foregoing own profit nor make themselves better off without inducing a benefit loss to the patient. We are specifically interested in the following questions: Does behaviour with regard to Pareto efficiency and tradeoffs vary in the two payment systems? Can a classification of behaviour help us to get deeper insights into decision making like it has helped to explain behaviour in other game settings (e.g. Selten and Ockenfels 1998 and Fischbacher et al. 2001)?

4 Experimental design and procedure

4.1 Design and parameters

We analyse physicians' provision behaviour under the two payment systems FFS and CAP. No other experimental parameter is varied. The experimental design allows for a controlled *ceteris paribus* variation and a between-subject comparison.

Each subject taking part in our experiment is allocated to a physician's role deciding on the quantity of medical services to be provided for given patients. Participants are medical students expected to become physicians in the future. We deliberately chose medical students as they most likely will identify with the decision task in our experiment. And we used a context-specific framing (see the instructions in the appendix). Both features are important as we are interested in how subjects decide in a medical context, and identification as well as framing seems to matter for behaviour.[10]

We run two treatments. In FFS, physicians receive a fee for each unit of medical service provided. In CAP, they are paid a lump-sum payment (capitation) per patient independent of the number of medical services they dispense. All monetary amounts are measured in Taler, our experimental currency, the exchange rate being 1 Taler = 0.05 EUR (about $0.07).

Our experimental physicians' task is to treat patients by providing them with medical services. Patients gain a benefit from these services. The patient benefit is measured in monetary terms. Three types of patients exist. These types differ in the 'benefit functions' that relate the benefit a patient receives to the number of services a physician provides. In particular, patients of different types need different amounts of services in order to get their optimal treatment (maximum benefit); for specifications of all functions see below. Patients in our experiment are abstract in that no real persons participate. Yet, to provide experimental physicians with an incentive for favourable behaviour towards the patients, the money corresponding to the benefits of all abstract patients is transferred to a charity caring for real patients.

Patients are further characterized by illnesses. An illness has no impact on patients' benefits. In FFS, it has an impact on physicians' remuneration, however, as the 'remuneration function' that relates a physician's remuneration to the number of services a physician provides is determined by the respective illness. In particular, maximum remunerations differ across the five existing illnesses. The same holds for maximum profits because the costs a physician has to bear are kept constant for all decisions and across treatments. Recall that in CAP, physicians are paid a lump-sum capitation per patient. Therefore, neither illnesses nor the number of medical services they dispense have an impact on their remuneration.

In the remainder of this subsection we describe the experimental design in more detail. Physicians decide on the quantity $q \in \{0, 1, \ldots, 10\}$ of medical services to be

[10] Ahlert et al. (2008) find less selfishness (higher identification) of physicians in a medical (familiar) framing than in a neutral (unfamiliar) environment. A recent own study shows non-medical students to behave much more selfish than medical students (Hennig-Schmidt and Wiesen 2009).

provided to their patients.[11] They decide for five abstract illnesses A, B, C, D, E[12] of three patient types 1, 2, 3. Patient types differ in their benefit from medical services ($B_1(q), B_2(q), B_3(q)$). Each combination of patient type and illness represents a specific patient 1A, 1B, 1C, ..., 3D, 3E (Table 1). By each decision ($j = 1, \dots, 15$), physicians simultaneously determine their own profit and the benefit of a given patient. The patient is assumed to be passive and fully insured accepting each medical service chosen by a physician.

Table 1: Order of decisions

Decision (j)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Patient type	1	1	1	1	1	2	2	2	2	2	3	3	3	3	3
Illness	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
Patient	1A	1B	1C	1D	1E	2A	2B	2C	2D	2E	3A	3B	3C	3D	3E

Physicians' remuneration. In FFS, physicians receive a fee for each unit of medical service provided. Fees differ across services and illnesses. As points of reference for our experimental fees we used tariffs for ophthalmologist services (like the treatment of glaucoma or cataract) taken from the German scale of charges and fees for physician services (EBM)[13]. Remuneration $R(q)$ increases in the quantity of medical services chosen (see Table 2).

In CAP, physicians are paid a lump-sum payment R per patient independent of their quantity decision. To make treatments comparable, R was specified at 12 Taler in CAP which is slightly above the average maximum profit per patient a physician could achieve in FFS (11.08 Taler).

Table 2: Physicians' remuneration $R(q)$

		Quantity (q)										
	Illness	0	1	2	3	4	5	6	7	8	9	10
FFS	A	0.00	1.70	3.40	5.10	5.80‡	10.50	11.00	12.10	13.50	14.90	16.60
FFS	B	0.00	1.00	2.40	3.50	8.00	8.40	9.40	16.00	18.00	20.00	22.50
FFS	C	0.00	1.80	3.60	5.40	7.20	9.00	10.80	12.60	14.40	16.20	18.30
FFS	D	0.00	2.00	4.00	6.00	8.00	8.20	15.00	16.90	18.90	21.30	23.60
FFS	E	0.00	1.00	2.00	6.00	6.70	7.60	11.00	12.30	18.00	20.50	23.00
CAP		12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00

‡ *Due to a display error on subjects' screens, physicians' remuneration for illness A at* $q_j = 4$ *was specified at 8.40 instead of 5.80. Physician's profits were displayed correctly, however. See the paragraph on physician's profit below.*

Patient benefit. Patients gain a benefit from medical services, the patient benefit $B(q)$ measured in monetary terms. Patient benefits vary across patient types. This reflects the heterogeneity of the patient population treated by a physician in reality, e.g. with regard to a patient's health status or different severities of illness. Table 3 shows patient benefits $B(q)$ given the quantity of medical services provided. A common characteristic of $B(q)$ is

[11] The range of services physicians can choose from may be interpreted as those eligible for a patient contracting with a certain health plan.

[12] We did not specify real illnesses because this turned out not to be feasible in the experimental setup.

[13] The German EBM lists medical services and the respective fees.

a global optimum $q^* \in [0, 10]$. The patient optimal quantity (q^*) yields the highest benefit $B(q^*_j)$ from medical services to the patient. The patient's optimal quantity is $q^*_j = 5$ for patient type 1 ($j = 1, \ldots, 5$), $q^*_j = 3$ for patient type 2 ($j = 6, \ldots, 10$) and $q^*_j = 7$ for patient type 3 ($j = 11, \ldots, 15$). After having reached the optimum, $B(q)$ declines because providing more medical services than q^* contributes negatively to a patient's benefit at the margin. Taking q^* as the benchmark for the right (best) medical treatment, we can identify overprovision and underprovision, respectively.

Table 3: Patient benefit $B(q)$

	Quantity (q)										
Patient type	0	1	2	3	4	5	6	7	8	9	10
1	0.00	0.75	1.50	2.00	7.00	10.00‡	9.50	9.00	8.50	8.00	7.50
2	0.00	1.00	1.50	10.00‡	9.50	9.00	8.50	8.00	7.50	7.00	6.50
3	0.00	0.75	2.20	4.05	6.00	7.75	9.00	9.45‡	8.80	6.75	3.00

‡ *Patient optimal quantity q^*_j yields the highest benefit $B(q^*_j)$ from medical services to the patient.*

It is crucial that the experimental physicians have an incentive to take the patient benefit into account. Therefore, the money corresponding to the benefits of all abstract patients aggregated over all decisions of all physicians was transferred to a charity caring for real patients – the *Christoffel Blindenmission*. To verify that the money was actually transferred we applied a procedure similar to the one used in Eckel and Grossman (1996). In each session, a monitor randomly selected from the participating subjects must verify, by a signed statement, that a check for the total patient benefit is written and sealed in an envelope addressed to the charity. The monitor and experimenter then walk together to the nearest mailbox and deposit the envelope. The monitor was paid an additional 4 EUR.

Physicians' profit. Further parameters relevant for physicians' decisions are costs and profit. Like real doctors, the experimental physicians have to bear costs depending on the quantity of medical services they choose. We use a convex cost function as assumed in several theoretical models (e.g. Ma 1994, 2007 and Choné and Ma 2007). $c(q_j) = 0.1q_j^2$ $\forall q \in [0, 10], j = 1, 2, \ldots, 15$ is applied in both treatments.

Profit (remuneration minus costs) varies across illnesses in FFS because fees differ for

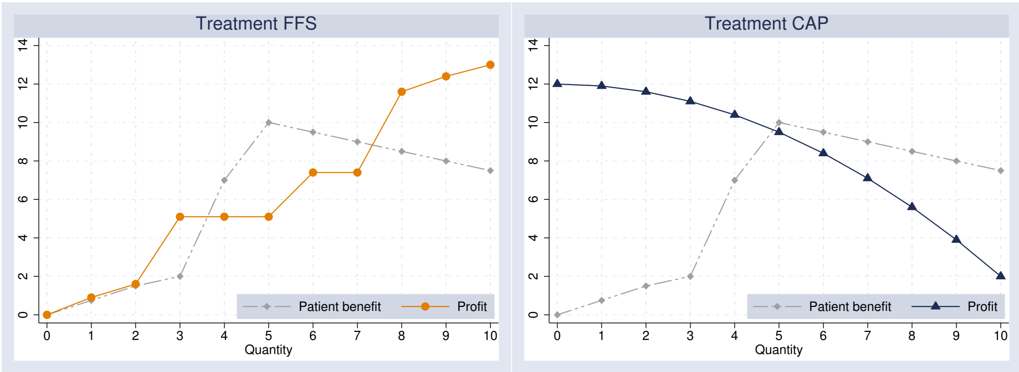
Table 4: Physicians' profit $\pi(q)$

		Quantity (q)										
	Illness	0	1	2	3	4	5	6	7	8	9	10
FFS	A	0.00	1.60	3.00	4.20	4.20	8.00‡	7.40	7.20	7.10	6.80	6.60
	B	0.00	0.90	2.00	2.60	6.40	5.90	5.80	11.10	11.60	11.90	12.50‡
	C	0.00	1.70	3.20	4.50	5.60	6.50	7.20	7.70	8.00	8.10	8.30‡
	D	0.00	1.90	3.60	5.10	6.40	5.50	11.40	12.00	12.50	13.20	13.60‡
	E	0.00	0.90	1.60	5.10	5.10	5.10	7.40	7.40	11.60	12.40	13.00‡
CAP		12.00‡	11.90	11.60	11.10	10.40	9.50	8.40	7.10	5.60	3.90	2.00

‡ *Physicians' maximum profit $\pi(\hat{q}_j)$ according to the profit-maximizing quantity of medical services $\hat{q}_j$.*

illnesses, and costs are the same for all patients. In CAP, on the other hand, profit does not vary with illnesses and patient types (see Table 4). Remember that patient benefit

Figure 1: Patient benefit and physician's profit for patient 1E (decision $j = 5$)



does, however.

For all patients in FFS, except for patient 1A ($j = 1$), the physician encounters a tradeoff between patient optimum and own profit maximization in that q_j^* differs from the profit maximizing quantity ($\hat{q}_j$). At $j = 1$ (patient 1A), $\hat{q}_j = q_j^* = 5$. For patient 3A ($j = 11$), $5 = \hat{q}_j < q_j^* = 7$. Except for illness A ($j = 1, 6, 11$) where $\hat{q}_j = 5$, the maximal profit is achieved at $q_j = 10$ (see left panel of Figure 1 for $j = 5$).

In CAP, $\hat{q}_j = 0$ for each decision $j = 1, \ldots, 15$. A higher patient benefit can only be achieved by a physician's deviating from her own maximal profit (see right panel of Figure 1 for $j = 5$).

4.2 Procedure

The computerized experiment was conducted in *B*onn*E*con*L*ab, the Laboratory for Experimental Economics at the University of Bonn. 42 medical students participated, 20 in FFS (one session) and 22 in CAP (two sessions). We thus base our analysis on 42 independent observations. Subjects were recruited by the online recruiting system ORSEE Greiner (2004) promising a monetary reward for participation in a decision-making task. The experiment was programmed using the software z-Tree (Fischbacher 2007).

Upon arrival, participants were randomly allocated to cubicles where they took their decisions in complete anonymity. Then, subjects were provided with the instructions that were read out aloud by the experimenter. Subjects was given plenty of time for clarifying questions which were asked and answered in private. To check for subjects' understanding of the experiment we asked them to answer three test questions structured like the actual experiment but with different parameter values. The experiment was not started unless all participants had answered all test questions correctly.

The experimental physicians then made their 15 quantity decisions the sequence of which was predetermined and kept across treatments (see Table 1). Having made their choices, subjects were asked to fill in a computerized questionnaire explaining their motivations and the factors having influenced their decisions. Finally, the monitor was assigned randomly. After the experiment, subjects were paid in private according to their choices. At last, the monitor verified that a check on the benefits of all patients was written and sealed in an envelope addressed to the *Christoffel Blindenmission*. The monitor and experimenter then walked together to the nearest mailbox and deposited the envelope.

Sessions lasted for about 40 minutes. The exchange rate per Taler was 0.05 EUR. On average subjects earned 6.88 EUR in FFS and 7.42 EUR in CAP.[14] In total, 273.68 EUR were transferred to the *Christoffel Blindenmission*, 6.62 EUR per participant in FFS and 6.42 EUR in CAP. The money supported surgical treatments of cataract patients in a hospital in Masvingo (Zimbabwe) staffed by ophthalmologists of the *Christoffel Blindenmission*. Average costs for such an operation amounted to 30 EUR. Thus, the money from our experiment allowed to treat nine patients. Note that subjects were not informed about the money being assigned to a developing country (see the instructions in the Appendix).

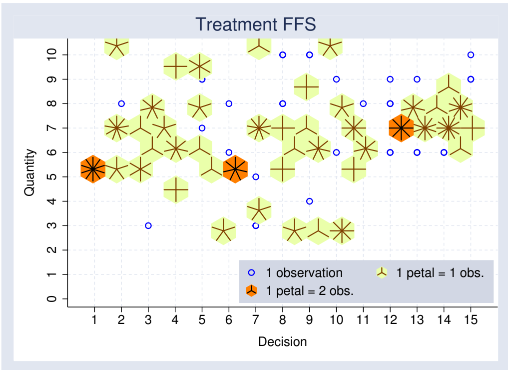
5 Results

In the present section, we investigate physicians' behaviour, both from the physician's and from the patient's perspective for FFS as well as for CAP. Moreover, we analyse the influence of physicians' profits and the patient benefit, and we study the impact of the payment system on patients' health status. We compare behaviour across treatments and, finally, we analyse how physicians' behaviour is affected by Pareto efficiency, i.e. by tradeoffs between physicians' profit and patient benefit.

5.1 Physicians' behaviour in FFS

Our first research question is related to behaviour under FFS. Remember that $\hat{q}_j = q_j^*$ for $j = 1$ (patient 1A), and $\hat{q}_j < q_j^*$ for $j = 11$ (patient 3A). Figure 3 shows absolute frequencies of all physicians' decisions for all patients. On average, 6.60 medical services are provided (median 7.00, SD 1.85). To study how patients are treated we analyse the quantity of

Figure 2: Absolute frequencies of quantity decisions per patient in FFS



medical services provided for each patient averaged over all physicians ($\bar{q}_j = \sum_{i=1}^{20} q_{ij}/20$).

Result FFS1. *In FFS, patients are overserved.*

[14] Average payoffs correspond to the hourly wage of a student helper at the University of Bonn (8.32 EUR). A lunch at the student cafeteria is around 2.50 EUR.

Support: $\overline{q}_j > q_j^*$ for the 13 patients where $\hat{q}_j > q_j^*$. Patient 1A ($j = 1$) is treated optimally by all physicians i, whereas patient 3A ($j = 11$) is underserved. Testing over all patients, $\overline{q}_j$ is highly significantly larger than q_j^* ($p = 0.002$, Wilcoxon signed ranks test, two-sided). Individual physicians largely deviate from choosing the patient optimal quantities. The mean deviation from q_j^*, $\overline{\mu}_i = \sum_{j=1}^{15}(q_{ij} - q_j^*)/15$, is positive for 17 of the 20 physicians, and zero for the remainder (see Table A.1 in the Appendix). Thus, significantly more physicians provide medical services that are larger than q_j^* ($p = 0.003$, binomial test, two-sided).

Next we investigate the impact of patient types on physicians' provision behaviour.

Result FFS2. *Overprovision in FFS depends on patient types.*

Support: Support is provided by test statistics of an order test (see Selten 1967 and Kuon, 1994) comparing the given order of average services per patient type with the perfect order (2, 1, 3) that accounts for q^* of each patient type.[15] There are six different possibilities to assign three ranks. The null hypothesis of the order test is that for each subject the order of observed values is arbitrary implying the mean inversion (standard deviation) being $\mu = 1.5$ ($\sigma = 0.9574$). As we observe 0.563 average inversions only, the null hypothesis can be rejected at the 1% level. A more in-depth analysis shows all patients of type 1 and 2 to be overserved (except for patient 1A) in that the number of physicians choosing $q_{ij} > q_j^*$ is larger than the number of physicians choosing $q_{ij} \leq q_j^*$. This is significant for four patients of type 1 and type 2 each ($p \leq 0.041$ binomial test, two-sided; see line I/FFS in Table A.2). Patients of type 3 are treated in a less consistent way. Patient 3A (3E) is underprovided (overprovided) and the remaining patients are treated optimally by at least half of the physicians.

Physician's profit. A physician's quantity decision determines her own profit. According to our research questions we are interested in whether profit maximizing is a main objective in general. As only 12% of the overall choices coincide with $\hat{q}_j$ this is rather not the case. The maximum profit $\pi(\hat{q}_j)$ a physician can achieve in FFS is 8.00 (12.50, 8.20, 13.60, 13.00) Taler for illness A (B, C, D, E); recall parameter values from Table 4. Choosing $\hat{q}_j$ for all j would have yielded an average payoff of 11.08 Taler. Physicians' actual quantity decisions resulted in an average overall profit of 9.17 Taler (median 8.00 Taler, SD 2.69 Taler), i.e. 17% lower than $\overline{\pi}(\hat{q}_j)$. Average profits for each physician i vary between 6.53 and 10.93 Taler (see Table A.4).[16] Testing over all patients, $\overline{\pi}(q_j)$ is highly significantly lower than $\overline{\pi}(\hat{q}_j)$ ($p = 0.001$, Wilcoxon signed ranks test, two-sided).

[15] The logic behind the order test is the following. When a physician's quantity choice is influenced by patient types (q^* per type), patients in need of a large (low) quantity of medical services should on average receive a large (low) amount of medical treatment. If a physician behaves accordingly the ranks assigned to the mean quantities provided per patient type should follow a "perfect order", namely 2, 1, 3. A measure for the difference between the actual order and the perfect order is the number of inversions, i.e. the number of pairwise changes necessary to transform the given order into the perfect order. We calculate the average quantity per patient type for each of those 16 physicians whose observed order comprises three different values and rank them according to their magnitude (see Table A.5). For each physician, we then calculate the number of inversions necessary to achieve the perfect order of ranks.

[16] For average profit per patient see Table A.3.

We are also interested in whether profits are affected by patient types. To this end, we study the deviation of each individual physician's profit from her profit maximum, i.e. $\hat{\pi}_j - \pi_{ij}$, for patient types separately. For the sake of comparability between FFS and CAP data, we compute for each patient the *relative* deviation $\Delta\pi_{ij} = (\hat{\pi}_j - \pi_{ij})/\hat{\pi}_j$. Table A.3 shows $\Delta\pi_{ij}$ averaged over all physicians. Highest deviations of up to 29% are found for patients 2B and 2E, whereas lowest deviations of less than 10% occur for patients 3A and 3C. There is no deviation for patient 1A because here all physicians choose their profit maximum that coincides with the patient benefit optimum. Average profit deviation is 14.66% for patients of type 1[17], 21.92% for those of type 2 and 11.98% for patients of type 3.

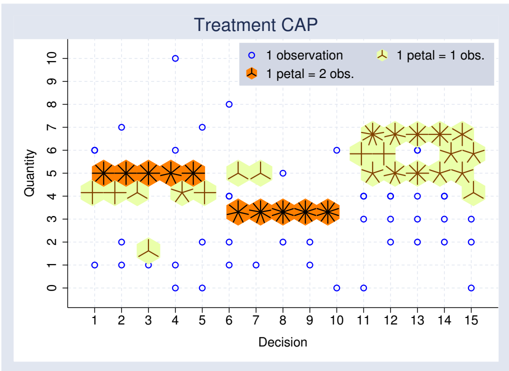
Patient benefit. A physician's decision also determines the patient benefit. In FFS – like in CAP – the benefit maximum for patients of type 3 ($B_3(q_j^*)$) is 9.45 Taler. $B_1(q_j^*) = B_2(q_j^*) = 10$ Taler (see Table 3). If physicians always chose the patient optimal quantity, patients would have received an average benefit $\overline{B}(q_j^*)$ of 9.82 Taler. Actual average patient benefit is 8.83 Taler (median 9.00 Taler, SD 1.10 Taler), i.e. 10% lower than $\overline{B}(q_j^*)$. Further, average patient benefits determined by physician i vary between 7.52 and 9.82 Taler (see Table A.4).

SUMMARY: Under FFS patients are overserved in that subjects on average choose quantities of medical services larger than the patient's optimal quantity. Provision is dependent on patient types as is the deviation of profits from the profit maximum. The levels of overprovision and of profit deviations tend to decrease with increasing needs of services. Physicians do not go for the maximal profit. This behaviour resulted in patients receiving a substantial benefit, only 10% on average less than the maximal amount.

5.2 Physicians' behaviour in CAP

Our second research question deals with behaviour under CAP. Recall that $0 = \hat{q}_j < q_j^*$ for all patients (decisions j). Figure 3 shows absolute frequencies of all physicians' decisions for all patients. On average, physicians chose 4.40 medical services (median 5.00, SD 1.64).

Figure 3: Absolute frequencies of quantity decisions per patient in CAP



[17]If we neglect patient 1A, $\Delta\pi_{ij} = 18.32\%$.

Result CAP1. *In CAP, patients are underserved.*

SUPPORT: $\overline{q}_j \leq q^*_j$ for 11 patients. Three patients (2A, 2B, 2C) are slightly overserved on average. Only patient 2E receives an optimal treatment on average. Testing over all patients, $\overline{q}_j$ is significantly smaller than q^*_j (p=0.0105, Wilcoxon signed ranks test, two-sided). Individual physicians largely deviate from the patient optimal quantity; but in contrast to FFS, they underserve in CAP. $\overline{\mu}_i$ is negative for 16 of the 22 physicians; $\overline{\mu}_i \geq 0$ for the remainder (Table A.1). Thus, weakly significanly more physicians choose quantites smaller than q^*_j ($p = 0.052$, Binomial test, two-sided).

Next we investigate whether underprovision is related to patient types.

Result CAP2. *Underprovision in CAP depends on patient types.*

SUPPORT: We again apply the order test and include those 19 subjects whose observed order comprises three different values (see Table A.5). Also in CAP, the order test reveals choices to be heavily dependent on patient types. We observe 0.158 average inversions. Thus the null hypothesis can be rejected at the 1% level. Analyzing the data in more detail shows that although patients are underserved on average, the number of physicians choosing q^*_j is larger than the number of physicians not choosing q^*_j for all patients of type 1 and 2. This is significant for four patients of type 2 (binomial test two-sided; see line I/CAP in Table A.2).[18] Patients of type 3 are underserved in that the number of physicians choosing $q_j < q^*_j$ is larger than the number of physicians choosing q^*_j.[19] This is weakly significant for one patient of type 3 (binomial test two-sided; see line I/CAP in Table A.2). Moreover, the level of underprovision $\overline{\nu}_j$ is highest for patient type 3 and lowest for patient type 2 (see Table A.3).

Physician's profit. The maximum profit $\pi(\hat{q}_j)$ a physician can achieve in CAP is 12.00 Taler for all illnesses (see Table 4). Physicians' actual quantity decisions resulted in an average profit $\overline{\pi}(q_j)$ of 9.79 Taler (median 9.50 Taler, SD 1.52 Taler), i.e. 18% lower than $\overline{\pi}(\hat{q}_j)$. Average profits for each physician i vary between 7.84 and 11.48 Taler (see Table A.4).[20] Testing over all patients, $\overline{\pi}(q_j)$ is highly significantly lower than $\overline{\pi}(\hat{q}_j)$ ($p = 0.000$, Wilcoxon signed ranks test, two-sided).

How are profits affected by patient types in CAP? Table A.3 shows $\Delta\pi_{ij}$ averaged over all physicians. Highest deviations of 25 to 30% are found for patients of type 3 whereas lowest deviations of 7 to 11% occur for patients of type 2. Average profit deviations are 18.75% (8.71%) for patients of type 1 (2) and 27.67% for those of type 3.

Patient benefit. The maximal average benefit a patient could gain in CAP, like in FFS, is 9.82 Taler if physicians always provided the patient optimal quantity. Actual average patient benefit is 8.56 Taler (median 9.75 Taler, SD 2.46 Taler), i.e. 13% lower than $\overline{B}(q^*_j)$.

[18] On average, 14.6 (18.4) physicians treat patients of type 1 (2) optimally, 6.2 (1.4) underprovide and 1.2 (2.2) overprovide.

[19] On average, 14.2 physicians underserve patients of type 3, 0.2 overprovides and 7.6 treat their patients optimally.

[20] For average profit per patient see Table A.3.

Further, average patient benefits determined by physician i vary between 2.73 and 9.82 Taler (see Table A.4).

SUMMARY: Under CAP patients are underprovided in that physicians on average choose quantities of medical services smaller than the patient's optimal quantity. Provision of services and the deviation of profits from the profit maximum are strongly influenced by patient types, i.e. with increasing needs for services the levels of underprovision and profit deviations tend to increase. Also in CAP, physicians do not strive for the maximal profit. Patients received a benefit being on average 13% lower than the maximum benefit.

5.3 Comparison of behaviour between FFS and CAP

Our third research question is related to the consequences of *both* payment systems. We are concerned with differences in the experimental physicians' behaviour across treatments and how patient types are affected. We compare physicians' profits, the provision of medical services, deviations from q_j^*, and patient benefit losses across payment systems for all patients and for patient types separately.

The results above have already shown that experimental physicians choose more medical services in FFS than in CAP. Thus, the next result implicitly follows from Results FFS1 and CAP1.

Result COMP1. *Patients are provided with more medical services in FFS than in CAP.*

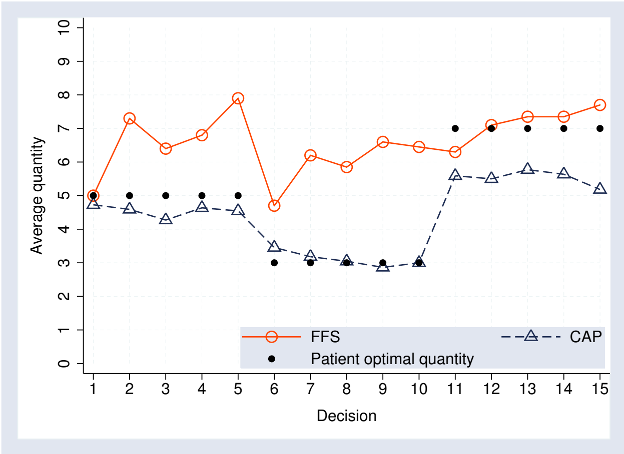
SUPPORT: Evidence is provided by Figure 4 showing the average quantity of medical services per decision (patient) in both treatments. Not only do physicians in FFS on average provide 50% more services than in CAP (6.60 vs. 4.40; median: 7.00 vs. 5.00; SD: 1.85 vs. 1.64) but for each decision j, $\overline{q}_j^{FFS} > \overline{q}_j^{CAP}$. This is highly significant ($p = 0.0000$, Mann-Whitney U test, two-sided). The picture is similar when comparing individual decisions across treatments for each patient. Except for patients 1A and 3A, q_{ij}^{FFS} is significantly larger than q_{ij}^{CAP} ($p \leq 0.0010$, Mann-Whitney U test, two-sided; see line II in Table A.2). Thus, in FFS a significantly higher number of patients is provided with significantly more medical services compared to CAP ($p = 0.007$, binomial test, two-sided).

Physician's profit. Physician's own profit $\pi(q_{ij})$ certainly is an important behavioural determinant in both treatments. As already mentioned, choosing $\hat{q}_j$ for all j in FFS would have yielded an average payoff $\overline{\pi}(\hat{q}_j)$ of 11.08 Taler. In CAP, the maximal profit is 12.00 Taler for all illnesses.

What did physicians actually do? They provided quantities of medical services such that their average profits are very similar in both treatments but about 17% lower than $\overline{\pi}(\hat{q}_j)$ (FFS: 9.17 Taler, CAP: 9.79 Taler). Average profits for each physician i vary between 6.53 and 10.93 Taler in FFS and between 7.84 and 11.48 Taler in CAP. In both payment systems, the average physician does not aim at the maximal achievable profit even though single physicians come very close to $\overline{\pi}(\hat{q}_j)$ (see Table A.4).

To answer the question how profits are affected by patient types, we compare $\Delta\pi_{ij}$

Figure 4: Average quantity of medical services per decision (patient)



across treatments. Except for patients 1C, 1E, 2C, and 3E, we find (weakly) significant differences between treatments.[21] For patients of type 2, $\Delta\overline{\pi}_{ij}^{FFS} > \Delta\overline{\pi}_{ij}^{CAP}$, for patients of type 3 the reverse holds.

Patient benefit and patient benefit loss. We next compare the two payment systems with regard to how patients' health status is affected by physicians' choices. To this end, we first focus on the optimal treatment and deviations thereof. We then concentrate on the benefit losses patients suffer on average when some or all of them are not treated optimally.

Result COMP2. *Patient optimal quantities exert a stronger influence on physicians' behaviour in CAP than in FFS.*

Support: Support comes from analyzing physicians' choices with regard to the patient optimal quantity. In CAP, the percentage of physicians choosing q_j^* per patient is significantly higher than in FFS ($p = 0.014$, Mann-Whitney U test, two-sided). If physicians deviate they tend towards opposite directions; a significantly larger share provides services larger than q_j^* in FFS compared to CAP ($p = 0.000$, Fisher exact test, two-sided). In FFS, $\overline{\mu}_i > 0$ except for physicians $i = 3, 4, 17$; in CAP, $\overline{\mu}_i \leq 0$ except for physicians $i = 4, 19$ (see Table A.1). Analyzing patient types separately, we find all patients of type 2 in CAP to get a better treatment in that significantly more physicians per patient chose q_j^* compared to FFS ($p \leq 0.011$, Fisher exact test; see line III in Table A.2). The same applies to all patients of type 1 except for patient 1A[22] ($p \leq 0.009$, Fisher exact test). Evidence is mixed for patients of type 3. We find no significant difference for patients 3A, 3C, 3E. For patients 3B and 3D physicians choose q_j^* significantly more often in FFS than in CAP (see line III in Table A.2). In both treatments, the benefit maximum for patients of type 3 ($B_3(q_j^*)$) is 9.45 Taler. $B_1(q_j^*) = B_2(q_j^*) = 10$ Taler resulting in $\overline{B}(q_j^*) = 9.82$ in FFS *and* in CAP. Our experimental physicians actually provide quantities of medical services such

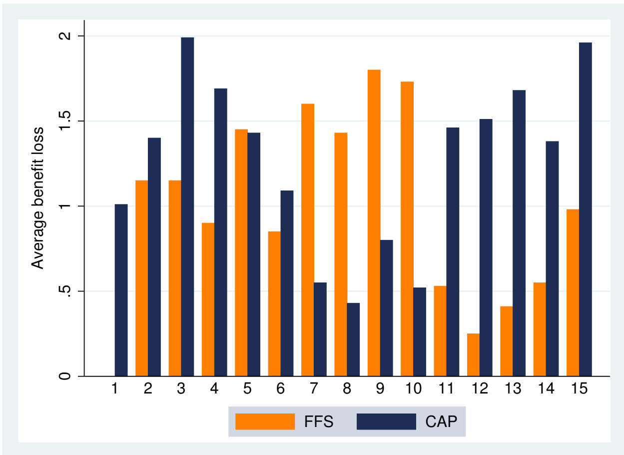
[21] For type 1: $p \leq 0.059$; for type 2: $p \leq 0.018$; for type 3: $p = 0.000$, all Mann-Whitney U test, two-sided; see line IV in Table A.2.

[22] Here, physicians in FFS make significantly more q_j^*-choices ($p = 0.006$, Fisher exact test; see line III in Table A.2).

that the average patient benefit $\overline{B}(q_j)$ was slightly larger in FFS (8.83 Taler) than in CAP (8.56 Taler) and around 10% smaller than $\overline{B}(q_j^*)$. These numbers seem to suggest that nearly no differences between payment systems exist. Yet, the picture is different when having a closer look at the data. Focusing on single patients and their health status we, like Newhouse and the Insurance Experiment Group (1993), find patients to be affected differently by the mode of payment (see below). Moreover, average patient benefits vary between 7.52 and 9.82 Taler in FFS and between 2.73 and 9.82 Taler in CAP (see Table A.4).

Whenever a physician i deviates from choosing q_j^* – when patients are either under- or overprovided – patients suffer a benefit loss ($\psi(q_{ji}) = |B(q_{ij}) - B(q_j^*)|$).

Figure 5: Average benefit loss per patient



Result COMP3. *Benefit losses per patient depend on patient types and differ across treatments.*

Support: Figure 5 contrasts the average benefit loss per patient across treatments. For 10 of the 15 patients, $\overline{\psi}(q_j)^{CAP} > \overline{\psi}(q_j)^{FFS}$. The benefits loss in FFS is larger for the remaining patients (see also Table A.3). Benefit losses differ significantly for all patients of type 2 ($p \leq 0.027$, Mann-Whitney U test, two-sided, see line IV in Table A.2). Losses are larger in FFS for patients 2B, ..., 2E; the reverse holds for patient 2A. For 9 of the 10 patients of types 1 and 3, benefit losses in CAP are larger than in FFS. Differences are only significant for two patients of type 1 and 3 each.[23]

Result COMP3 suggests that for patients in need of a small quantity of medical services like patients of type 2, a smaller benefit loss results when physicians are paid by CAP. Patients in need of a larger quantity of medical services, like patients of types 1 and 3, incur a smaller loss under FFS.

Summary: The cross-treatment comparison demonstrates that physicians' choices are highly influenced by the payment system. Physicians in FFS choose more medical services

[23] 1A (where no losses occur in FFS): $p = 0.009$; 1E: $p = 0.062$; 3B: $p = 0.002$; 3C: $p = 0.050$, all Mann-Whitney U test, two-sided, see line IV in Table A.2.

than those in CAP do. Consequently, the mode of payment does affect patients' health status. In particular, patients of type 1 and 2 are treated more optimally under CAP than under FFS and the patient benefit loss is significantly smaller in the former payment system for all but one of patients of type 2.

5.4 Tradeoffs and Pareto efficiency

In this section, we are concerned with the tradeoff between own profit and patient benefit a physician encounters when making a quantity decision. In particular, we investigate how Pareto efficiency influences a physician's behaviour.

In general, Pareto efficiency means that an allocation X is Pareto preferred to another allocation Y if at least one person is better off and no one is worse off with X than with Y. Besides its importance in general economic theory, the concept of Pareto efficiency also plays a prominent role in health economics (e.g. Iversen 1993, De Jaegher and Jegers 2000 and Pau and Vera-Hernandez 2007). In the context of our experiment a situation is said to be Pareto efficient, if no unanimous move to another allocation of profit and patient benefit is possible. That means, a Pareto-efficient (PE) choice involves that changing q can neither make the physician better off without inducing a benefit loss to the patient nor make the patient better off without foregoing own profit. Pareto-inefficient (PIE) choices do not involve a benefit/profit tradeoff as changing q can increase *both* a physician's own profit and the patient benefit; they are dominated by Pareto-efficient choices.

Pareto-efficient decision options exist for each patient in both treatments. The number of PE benefit/profit pairs differs according to illnesses (in FFS only) and patient types. In FFS, physicians can choose between one and eight PE decisions per patient. In CAP, there are either four (patient type 2), six (patient type 1) or eight (patient type 3) PE pairs. PE choices are positioned on the upper right line in Figures A.1 and A.2, the Pareto frontier, whereas PIE decisions are those below the Pareto frontier.

It is remarkable that 597 of the 660 choices are Pareto-efficient. Thus, 95% of all physicians' actual choices both in FFS and CAP involve a tradeoff between physician's own profit and patient benefit. Pareto efficiency guides all the decisions by 13 of the 20 physicians (65%) in FFS and by 15 of the 22 physicians (68%) in CAP. The remaining choices entail up to 4 (9) PIE decisions per physician in FFS (CAP). Hence, not only has the majority of physicians Pareto efficiency as their only target but also the remaining physicians behave accordingly with the vast majority of their quantity decisions.

To further characterize physicians' choices we subdivide the set of PE decisions into categories capturing variables of economic importance and medical ethics: own profit maximum, patient benefit optimum, social optimum.

- PROMAX comprises choosing $\hat{q}_j$, the profit-maximizing quantity of medical services. The corresponding benefit/profit pair is $(B(\hat{q}_j), \pi(\hat{q}_j))$.
- PATMAX consists of q_j^*-choices maximizing the patient's benefit. $(B(q_j^*), \pi(q_j^*))$ is the resultant benefit/profit pair.

  PROMAX and PATMAX are the two boundary points of the Pareto frontier (see

Figures A.1 and A.2).

- SOCOPT is suggested by a welfare economics perspective and contains the socially optimal choices, i.e. decisions where $(\pi(q_j) + B(q_j))$ is maximal.
  Note that patients exist where SOCOPT coincides with PROMAX and/or PATMAX (Table A.6). Only those decisions are assigned to SOCOPT that are not yet covered by the two previous categories.[24]

- PAROTH is a residual category comprising the remaining benefit/profit pairs on the Pareto frontier not included in any of the other three categories.

In FFS, 16% of all physician's Pareto-efficient choices are assigned to PROMAX, 34% to PATMAX, 16% to SOCOPT and 34% to PAROTH.[25] In CAP, only 2% of physicians' choices are attributed to PROMAX, 66% to PATMAX, 6% to SOCOPT and 26% are covered by PAROTH. [26]

Comparing both payment systems, a much lower percentage of decisions in CAP is motivated by $\pi(\hat{q}_j)$ probably because choosing $\hat{q}_j^{CAP}$ entails *no* provision of services to the patient. Such behaviour would be a severe violation of the professional code of medical ethics. Noticeably, two thirds of all Pareto-efficient decisions in CAP involve $B(q_j^*)$ versus only one third in FFS. This may be due to the fact that choosing q^{*CAP} implies a lower own-profit reduction than in FFS where the physician on average forgoes 39.6% of her maximally achievable profit vs. only 23.3% in CAP. The social optimum plays no role in CAP possibly because q^{soc} coincides with q^* for all 10 patients of types 1 and 2.

SUMMARY: Our analysis provided compelling results. First, nearly all physicians' decisions are influenced by Pareto efficiency. Second, the vast majority of these choices (66% in FFS and 74% in CAP) can be explained by motives based on variables of economic and ethical importance.

6 Conclusion

The paper introduces a controlled laboratory experiment to test for the influence of payment systems on physicians' provision behaviour. By assigning the monetary equivalent of the patient benefit to treating actual patients we substituted the 'abstract' patients in our experiment with real ones. Our design was successful in eliciting benevolent behaviour towards the patient. Not only were mean benefit losses rather low (10% in FFS and 13% in CAP) but also did nearly all experimental physicians in a post-experimental questionnaire state the patient benefit to have influenced their decisions.

Our results are in line with the theoretical literature (e.g. Ellis and McGuire 1986)

[24] We decided on this assignment as subjects choices may not be motivated by the social optimum in the first place for the following reason. Finding q^{soc} seems not straightforward as participants first have to calculate $\pi(q_j) + B(q_j)$ and then they have to determine the maximum. Selecting $\hat{q}$ or q^* is much more obvious given the information on the decision screens (see the Instructions in the Appendix).

[25] When calculating the percentages, $j = 1$ is neglected because here categories PROMAX, PATMAX and SOCOPT coincide and we cannot even distinguish whether $q = 5$ was motivated by $\hat{q}_j$ or by q_j^*.

[26] A detailed overview on relative frequencies per category is provided in Table A.6.

and add further evidence to generalizing previous empirical findings in the field. Patients are overserved in FFS in that experimental physicians on average choose quantities of medical services larger than the patient's optimal quantity. Provision is dependent on patient types as is the deviation of profits from the profit maximum. The cross-treatment comparison most clearly shows physicians' choices to be highly influenced by the payment system. Physicians in FFS provide more medical services than those in CAP do. Like Newhouse and the Insurance Experiment Group (1993), we found the mode of payment to affect patients' health status. Patients in need of a low level of medical services are better off under CAP, whereas patients with a high need of medical services gain more health benefit when physicians are paid by FFS. How these gains and losses are to be weighed against each other is a matter of political decision, however.

In both remuneration systems, financial incentives are not the only motivation for physicians' quantity decisions, though. As the patient benefit is of considerable importance, patients received a substantial benefit the financial equivalent of which allowed to treat nine real patients by ophthalmic surgery.

Experiments in health economics might serve as a 'wind tunnel' or 'test bed' before institutional changes are implemented during a health care reform. Even though an experiment always simplifies a physician's decision task when caring for a patient it, at the same time, allows to separating behavioural determinants. While simplifications give rise to caution when extrapolating the results, they also suggest the lines for further experimental research like introducing uncertainty about the impact of medical treatments and patients' health status, patients' demand effects and monitoring mechanisms.

References

Ahlert, M., S. Felder, and B. Voigt (2008): "How Economists and Physicians Trade Off Efficiency and Equity in Medically and Neutrally Framed Allocation Problems?" FOR 655 Working Paper No. 09/2008, Jacobs University Bremen.

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

Barros, P. (2003): "Cream-Skimming, Incentives for Efficiency and Payment System," *Journal of Health Economics*, 22, 419–443.

Blomqvist, A. (1991): "The doctor as double agent: Information asymmetry, health insurance, and medical care," *Journal of Health Economics*, 10, 411–432.

Camerer, C. and R. Hogarth (1999): "The Effects of Financial Incentives in Economics Experiments: A Review and Capital-production-labor Framework," *Journal of Risk and Uncertainty*, 18, 7–42.

Choné, P. and C. Ma (2007): "Asymmetric Information from Physician Agency: Optimal Payment and Health Care Quantity," Working Paper, University of Boston.

Davidson, S., L. Manheim, S. Werner, M. Hohlen, B. Yudkowsky, and G. Flemming (1992): "Prepayment with Office-based Physicians in Publicly Funded Programs: Results from Children's Medicaid Program," *Pediatrics*, 89, 761–767.

Davis, D. and C. Holt (1993): *Experimental Economics*, Princeton University Press, Princeton, N.J.

De Jaegher, K. and M. Jegers (2000): "A Model of Physician Behaviour with Demand Inducement," *Journal of Health Economics*, 19, 231–258.

Devlina, R. and S. Sarma (2008): "Do Physician Remuneration Schemes Matter? The Case of Canadian Family Physicians," *Journal of Health Economics*, 27, 1168–1181.

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

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

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

Falk, A. and E. Fehr (2003): "Why Labor Market Experiments?" *Labor Economics*, 10, 399–406.

Fan, C., K. Chen, and K. Kan (1998): "The Design of Payment Systems for Physicians under Global Budget - An Experimental Study," *Journal of Economic Behavior and Organization*, 34, 295–311.

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

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71, 397–404.

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

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

Gosden, T., F. Forland, I. Kristiansen, M. Sutton, B. Leese, A. Guiffrida, M. Sergison, and L. Pedersen (2001): "Impact of Payment Method on Behavior of Primary Care Physicians: A Systematic Review," *Journal of Health Services Research and Policy*, 6, 44–54.

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

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

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

Hennig-Schmidt, H. and D. Wiesen (2009): "Are Medical Students more Caring in Incentive-based Decision Tasks? An Experimental Study on Payment System Variation," unpublished manuscript.

Hertwig, R. and A. Ortmann (2001): "Experimental Practices in Economics: A Methodological Challenge for Psychologists?" *Behavioral and Brain Sciences*, 24, 383–403.

Hutchinson, B., S. Birch, J. Hurley, J. Lomas, and F. Stratford-Devai (1996): "Do Physician-payment Mechanisms Affect Hospital Utilization? A Study of Health Service Organizations in Ontario," *Canadian Medical Association Journal*, 154, 653–661.

Iversen, T. (1993): "A Theory of Hospital Waiting Lists," *Journal of Health Economics*, 12, 55–71.

Iversen, T. and H. Lurås (2000): "The Effect of Capitation on GPs' Referal Decision," *Health Economics*, 9, 199–210.

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

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

Kuon, B. (1994): *Two-Person Bargaining Experiments with Incomplete Information*, vol. 412 of *Lecture Notes in Economics and Mathematical Systems*, Berlin, Heidelberg, New York: Springer Verlag.

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

——— (2007): "Altruism and Incentives in Public and Private Health Care," in *Finance and Incentives of the Health Care System, Proceedings of the 50th Anniversary Symposium of the Yrjö JahnssonFoundation*, ed. by A. Suvanto and H. Vartiainen, Government Institute for Economic Research, Helsinki, 79–104.

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

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

——— (2002): *Pricing the Priceless – A Health Care Conundrum*, Cambridge (MA), MIT Press.

Newhouse, J. and the Insurance Experiment Group (1993): *Free for All: Lessons from the RAND Health Insurance Experiment*, Harvard University Press.

Pau, O. and M. Vera-Hernandez (2007): "Competition among Differentiated Health Plans under Adverse Selection," *Journal of Health Economics*, 26, 233–250.

Selten, R. (1967): "Ein Oligopolexperiment mit Preisvariation und Investition," in *Beiträge zur experimentellen Wirtschaftsforschung*, ed. by H. Sauermann, J.C.B. Mohr (Paul Siebeck), Tübingen, 105–135.

Selten, R. and A. Ockenfels (1998): "An Experimental Solidarity Game," *Journal of Economic Behavior and Organization*, 34, 517–539.

Stearns, S., B. Wolfe, and D. Kindig (1992): "Physician Responses to Fee-for-service and Capitation Payment," *Inquiry*, 29, 416–425.

A Appendix

Instructions [*translated from German*]

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 to read the instructions carefully.

You take your decisions anonymously in your cubicle 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 Taler. At the end of the experiment, your earnings will be converted into Euro at an exchange rate of 1 Taler = 0.05 EUR and paid to you in cash.

Your decisions in the experiment

During the entire experiment you are in the role of a physician. You have to decide on the treatment of 15 patients. All participants of this experiment are taking their decisions in the role of a physician. You decide on the **quantity** of medical services you want to provide for a given illness of a patient.

You decide on your computer screen where five different illnesses – 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:

- *Treatment FFS*:
  A different PAYMENT is assigned to each **quantity** of medical services . The PAYMENT increases in the **quantity** of medical services.
- *Treatment CAP*:
  For each patient you receive a lump-sum PAYMENT that is independent of 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 Taler is calculated by subtracting your COSTS from your PAYMENT.

To each quantity of medical services a certain benefit for the patient is assigned, the PA-

TIENT 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.

Screen shot FFS

Illness E / Patient 1

Medical services	Quantity	Your Payment (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 E1	1	1.00	0.10	0.90	0.75
Service E1, Service E2	2	2.00	0.40	1.60	1.50
Service E1, Service E2, Service E3	3	6.00	0.90	5.10	2.00
Service E1, Service E2, Service E3, Service E4	4	6.70	1.60	5.10	7.00
Service E1, Service E2, Service E3, Service E4, Service E5	5	7.60	2.50	5.10	10.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6	6	11.00	3.60	7.40	9.50
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7	7	12.30	4.90	7.40	9.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8	8	18.00	6.40	11.60	8.50
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8, Service E9	9	20.50	8.10	12.40	8.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8, Service E9, Service E10	10	23.00	10.00	13.00	7.50

Please indicate the quantity of medical services you wish to provide:

OK

Screen shot CAP

Illness E / Patient 1

Medical services	Quantity	Your Payment (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 E1	1	12.00	0.10	11.90	0.75
Service E1, Service E2	2	12.00	0.40	11.60	1.50
Service E1, Service E2, Service E3	3	12.00	0.90	11.10	2.00
Service E1, Service E2, Service E3, Service E4	4	12.00	1.60	10.40	7.00
Service E1, Service E2, Service E3, Service E4, Service E5	5	12.00	2.50	9.50	10.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6	6	12.00	3.60	8.40	9.50
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7	7	12.00	4.90	7.10	9.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8	8	12.00	6.40	5.60	8.50
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8, Service E9	9	12.00	8.10	3.90	8.00
Service E1, Service E2, Service E3, Service E4, Service E5 Service E6, Service E7, Service E8, Service E9, Service E10	10	12.00	10.00	2.00	7.50

Please indicate the quantity of medical services you wish 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 named "Your Decision".

There are no real but abstract patients participating in this experiment. Yet the PATIENT

BENEFIT an abstract patient receives by your providing medical services will be beneficial for a real patient. The total amount corresponding to the sum over all 15 PATIENT BENEFITS determined by your decisions will be transferred to the charity *Christoffel Blindenmission Deutschland e.V., 64625 Bensheim*, to support an ophthalmic hospital where patients with cataract are treated.

Earnings in the experiment

After having made your 15 decisions, your overall earnings will be calculated by summing up the PROFITs from all your decisions. This amount will be converted from Taler into Euro at the end of the experiment.

The overall PATIENT BENEFIT resulting from your 15 quantity decisions will be converted into Euro as well and will be transferred to the *Christoffel Blindenmission*.

The transferral will be made by the experimenter and a monitor. The monitor writes a check on the amount of money corresponding to the aggregated PATIENT BENEFITS of this experiment. This check issued to the *Christoffel Blindenmission* will be sealed in an envelope addressed to this charity. The monitor and experimenter then walk together to the nearest mailbox and deposit the envelope.

After all participants have taken their decisions, one participant is randomly assigned the role of the monitor. The monitor receives a payment of 4 EUR in addition to the payment from the experiment. The monitor verifies, by a signed statement, that the procedure described above was actually carried out.

Next, please answer some questions familiarizing you with the decision situation.

After your 15 decisions, please answer some further questions on your screen.

Table A.1: Descriptive statistics on quantity (q_{ij}) and differences to patient optimal quantity (μ_i) per physician i

	FFS						CAP					
	q_{ij}			μ_i			q_{ij}			μ_i		
i	Mean ($\overline{q}_i$)	Median	SD	Mean ($\overline{\mu}_i$)	Median	SD	Mean ($\overline{q}_i$)	Median	SD	Mean ($\overline{\mu}_i$)	Median	SD
1	6.40	7.00	1.12	1.40	1.00	1.35	4.20	5.00	1.52	-0.80	-1.00	1.61
2	7.73	8.00	1.87	2.73	2.00	2.94	4.27	5.00	0.88	-0.73	0.00	1.28
3	5.00	5.00	1.46	0.00	0.00	0.65	4.80	5.00	1.47	-0.20	0.00	0.41
4	5.00	5.00	1.69	0.00	0.00	0.00	5.13	5.00	1.60	0.13	0.00	0.52
5	7.27	8.00	1.16	2.27	2.00	1.49	2.13	2.00	0.83	-2.87	-4.00	1.46
6	6.40	6.00	1.12	1.40	1.00	1.80	5.00	5.00	1.69	0.00	0.00	0.00
7	7.13	7.00	1.06	2.13	2.00	1.77	4.07	4.00	0.96	-0.93	-1.00	1.10
8	8.27	9.00	1.94	3.27	3.00	2.69	4.33	5.00	0.98	-0.67	0.00	0.98
9	6.07	7.00	1.39	1.07	1.00	1.28	4.07	4.00	0.80	-0.93	-1.00	1.22
10	7.67	7.00	1.76	2.67	2.00	2.55	5.00	5.00	1.69	0.00	0.00	0.00
11	7.47	8.00	2.00	2.47	2.00	2.61	4.93	5.00	1.62	-0.07	0.00	0.26
12	6.93	7.00	1.75	1.93	2.00	2.05	4.93	5.00	1.62	-0.07	0.00	0.26
13	6.13	6.00	1.92	1.13	1.00	1.92	2.40	2.00	1.18	-2.60	-2.00	2.38
14	6.27	7.00	1.33	1.27	1.00	1.62	5.00	5.00	1.69	0.00	0.00	0.00
15	8.53	9.00	1.96	3.53	3.00	2.85	4.00	4.00	0.85	-1.00	-1.00	0.85
16	6.67	6.00	1.54	1.67	1.00	2.47	4.47	5.00	1.85	-0.53	0.00	2.00
17	5.00	5.00	1.69	0.00	0.00	0.00	3.40	4.00	1.68	-1.60	-3.00	1.99
18	5.73	6.00	1.49	0.73	1.00	1.03	4.53	5.00	1.19	-0.47	0.00	0.74
19	7.00	7.00	1.25	2.00	2.00	1.96	6.00	6.00	2.45	1.00	2.00	2.56
20	5.33	5.00	1.45	0.33	0.00	1.11	4.67	5.00	1.29	-0.33	0.00	0.49
21							5.00	5.00	1.69	0.00	0.00	0.00
22							4.47	5.00	1.13	-0.53	0.00	0.83

Table A.2: Test statistics of two-sided non-parametric tests per patient

	Test; Variable(s); Scope	Decision j (Patient)														
		1 (1A)	2 (1B)	3 (1C)	4 (1D)	5 (1E)	6 (2A)	7 (2B)	8 (2C)	9 (2D)	10 (2E)	11 (3A)	12 (3B)	13 (3C)	14 (3D)	15 (3E)
I	Binomial; q_j^*, $\neg q_j^*$; FFS	0.0000	0.0414	0.1153	0.0118	0.0025	0.0414	0.0004	0.0414	0.0025	0.5034	0.5034	0.0414	1.0000	0.5034	0.0118
	Binomial; q_j^*, $\neg q_j^*$; CAP	0.1338	0.1338	0.1338	0.5235	0.1338	0.2863	0.0043	0.0001	0.0001	0.0001	0.5235	0.1338	0.2863	0.2863	0.0524
II	Mann Whitney U; q_{ij}; across treatments	0.2440	0.0000	0.0000	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000	0.0001	0.2339	0.0000	0.0002	0.0001	0.0000
III	Fisher exact; q_j^*; across treatments	0.0063	0.0051	0.0000	0.0095	0.0005	0.0111	0.0000	0.0000	0.0000	0.0006	0.2461	0.0051	0.1670	0.0784	0.2457
IV	Mann Whitney U; $\psi(q_{ij})$; across treatments	0.0066	0.1539	0.3047	0.4630	0.0617	0.0271	0.0000	0.0000	0.0000	0.0008	0.3972	0.0028	0.0507	0.1206	0.5991
V	Mann Whitney U; $\Delta\pi_{ij}$; across treatments	0.0000	0.0195	0.8055	0.0590	0.2046	0.0183	0.0008	0.2092	0.0119	0.0203	0.0000	0.0002	0.0000	0.0000	0.4219

Table A.3: Descriptive statistics on variables per patient

			Decision j (Patient)														
			1 (1A)	2 (1B)	3 (1C)	4 (1D)	5 (1E)	6 (2A)	7 (2B)	8 (2C)	9 (2D)	10 (2E)	11 (3A)	12 (3B)	13 (3C)	14 (3D)	15 (3E)
FFS	q_{ij}	Mean ($\overline{q}_j$)	5.00	7.30	6.40	6.80	7.90	4.70	6.20	5.85	6.60	6.45	6.30	7.10	7.35	7.35	7.70
		Median	5.00	7.00	6.50	6.80	8.00	5.00	7.00	6.00	6.00	8.00	6.00	7.00	7.00	7.00	8.00
	ν_{ij}	Mean ($\overline{\nu}_j$)	0.00	2.30	1.40	1.80	2.90	1.70	3.20	2.85	3.60	3.45	-0.70	0.10	0.35	0.35	0.70
		Median	0.00	2.00	1.50	1.00	3.00	2.00	4.00	3.00	3.00	5.00	-1.00	0.00	0.00	0.00	1.00
		SD	0.00	1.84	1.50	1.67	1.86	1.22	2.24	2.21	2.23	3.05	0.86	0.64	0.75	0.88	1.03
	$\pi(q_{ij})$	Mean	8.00	10.18	7.21	10.79	10.29	6.98	9.03	6.68	10.88	9.18	7.43	10.66	7.78	12.20	10.28
		Median	8.00	11.10	7.45	11.40	11.60	8.00	11.10	7.20	11.40	11.60	7.40	11.10	7.70	12.00	11.60
		SD	0.00	2.60	0.90	2.82	3.04	1.66	3.26	1.41	2.92	3.62	0.31	1.68	0.25	0.51	2.20
	$\Delta\pi_{ij}$	Mean	0.00	0.19	0.12	0.21	0.21	0.13	0.28	0.20	0.20	0.29	0.07	0.15	0.06	0.10	0.21
		Median	0.00	0.11	0.09	0.16	0.11	0.00	0.11	0.13	0.16	0.11	0.08	0.11	0.07	0.12	0.11
		SD	0.00	0.21	0.11	0.21	0.23	0.21	0.26	0.17	0.21	0.28	0.04	0.13	0.03	0.04	0.17
	$B(q_{ij})$	Mean	10.00	8.85	8.85	9.10	8.55	9.15	8.40	8.58	8.20	8.28	8.92	9.21	9.04	8.90	8.47
		Median	10.00	9.00	9.00	9.50	8.50	9.00	8.00	8.50	8.50	7.50	9.00	9.45	9.23	9.45	8.80
	$\psi(q_{ij})$	Mean ($\overline{\psi}_j$)	0.00	1.15	1.15	0.90	1.45	0.85	1.60	1.43	1.80	1.73	0.53	0.25	0.41	0.55	0.98
		Median	0.00	1.00	1.00	0.50	1.50	1.00	2.00	1.50	1.50	2.50	0.45	0.00	0.23	0.00	0.65
		SD	0.00	0.92	1.73	0.84	0.93	0.61	1.12	1.10	1.12	1.53	0.64	0.62	0.62	0.96	1.48
		N	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
CAP	q_{ij}	Mean ($\overline{q}_j$)	4.73	4.59	4.27	4.64	4.55	3.45	3.18	3.05	2.86	3.00	5.59	5.50	5.77	5.64	5.18
		Median	5.00	5.00	5.00	5.00	5.00	3.00	3.00	3.00	3.00	3.00	6.00	5.50	5.50	6.00	5.50
	ν_{ij}	Mean ($\overline{\nu}_j$)	-0.27	-0.41	-0.73	-0.36	-0.45	0.45	0.18	0.05	-0.14	0.00	-1.41	-1.50	-1.23	-1.36	-1.82
		Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-1.00	-1.50	-1.50	-1.00	-1.50
		SD	0.98	1.18	1.28	1.81	1.34	1.37	0.85	0.49	0.47	0.93	1.74	1.41	1.69	1.43	1.82
	π_{ij}	Mean	9.67	9.76	10.02	9.54	9.76	10.63	10.92	11.05	11.16	11.02	8.59	8.79	8.40	8.63	9.00
		Median	9.50	9.50	9.50	9.50	9.50	11.10	11.10	11.10	11.10	11.10	8.40	8.95	8.95	8.40	8.95
		SD	0.72	0.92	0.86	1.87	0.93	1.29	0.60	0.36	0.20	0.62	1.52	1.42	2.02	1.45	1.55
	$\Delta\pi_{ij}$	Mean	0.19	0.19	0.17	0.21	0.19	0.11	0.09	0.08	0.07	0.08	0.28	0.27	0.30	0.28	0.25
		Median	0.21	0.21	0.21	0.21	0.21	0.08	0.08	0.08	0.08	0.08	0.30	0.25	0.25	0.30	0.25
		SD	0.06	0.08	0.07	0.16	0.08	0.11	0.05	0.03	0.02	0.05	0.13	0.12	0.17	0.12	0.13
	$B(q_{ij})$	Mean	8.99	8.60	8.01	8.31	8.57	8.91	9.45	9.57	9.20	9.48	7.99	7.94	7.77	8.07	7.49
		Median	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	9.00	8.38	7.75	9.00	8.38
	$\psi(q_{ij})$	Mean ($\overline{\psi}_j$)	1.01	1.40	1.99	1.69	1.43	1.09	0.55	0.43	0.80	0.52	1.46	1.51	1.68	1.38	1.96
		Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.45	1.08	1.70	0.45	1.08
		SD	2.18	2.69	3.40	2.88	2.80	2.55	1.92	1.81	2.58	2.14	2.30	1.91	2.14	1.94	2.56
		N	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22

Table A.4: Descriptive statistics on patient benefit $B(q_{ij})$ and profit $\pi(q_{ij})$ per physician i

	FFS						CAP					
	Patient benefit $B(q_{ij})$			Profit $\pi(q_{ij})$			Patient benefit $B(q_{ij})$			Profit $\pi(q_{ij})$		
i	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
1	9.11	9.45	0.55	8.83	8.00	1.93	8.25	9.00	2.54	10.02	9.50	1.03
2	8.19	8.50	1.08	10.09	11.60	2.79	8.67	9.00	1.43	10.11	9.50	0.71
3	9.69	10.00	0.42	6.73	5.90	2.27	9.73	10.00	0.42	9.49	9.50	1.41
4	9.82	10.00	0.27	6.53	5.90	2.51	9.75	10.00	0.34	9.13	9.50	1.62
5	8.63	8.80	0.59	10.15	11.10	2.11	2.73	1.50	2.45	11.48	11.60	0.34
6	9.04	9.00	0.75	9.50	11.10	2.26	9.82	10.00	0.27	10.26	10.40	0.82
7	8.66	8.80	0.63	10.15	11.10	1.98	8.50	9.00	1.61	11.48	11.60	0.34
8	7.54	7.75	1.63	10.81	12.00	2.38	9.25	10.00	1.10	9.23	9.50	1.70
9	9.22	9.00	0.45	8.88	8.00	2.37	8.38	7.75	1.59	10.03	9.50	0.78
10	8.47	8.80	1.10	10.46	11.10	2.50	9.82	10.00	0.27	10.29	10.40	0.64
11	7.68	7.50	1.90	10.24	11.10	2.76	9.79	10.00	0.33	9.23	9.50	1.70
12	8.84	9.00	0.89	9.81	11.10	2.64	9.79	10.00	0.33	9.32	9.50	1.62
13	9.18	9.45	0.86	8.99	8.00	2.90	4.87	2.20	3.91	11.29	11.60	0.55
14	9.17	9.45	0.73	9.37	11.10	2.41	9.82	10.00	0.27	9.23	9.50	1.70
15	7.52	7.50	1.10	10.93	12.40	2.48	8.25	7.75	1.32	10.33	10.40	0.68
16	8.79	9.00	0.95	9.70	11.10	2.30	8.74	9.50	2.58	9.69	9.50	1.48
17	9.82	10.00	0.27	6.53	5.90	2.51	6.47	6.00	3.54	10.58	10.40	0.97
18	9.38	9.45	0.44	8.69	8.00	2.52	9.50	10.00	0.82	9.81	9.50	1.03
19	8.51	8.50	0.88	10.11	11.10	1.79	7.23	8.50	2.90	7.84	8.40	2.90
20	9.39	9.00	0.49	6.87	6.50	1.68	9.67	10.00	0.49	9.67	9.50	1.15
21							9.82	10.00	0.27	9.23	9.50	1.70
22							9.42	10.00	0.93	9.89	9.50	0.96

Table A.5: Order test

	Sub.	Average quantity pat. type 1	pat. type 2	pat. type 3	Order			Number of inversions
FFS	1	6.4	5.6	7.2	2	1	3	0
	2	8.0	8.4	6.8	2	3	1	1
	3	5.0	3.4	6.6	2	1	3	0
	4	5.0	3.0	7.0	2	1	3	0
	5	7.2	6.6	8.0	2	1	3	0
	6	5.8	6.4	7.0	1	2	3	1
	7	7.2	6.8	7.4	2	1	3	0
	8	8.4	8.4	8.0	2	2	1	-
	9	6.8	4.6	6.8	2	1	2	-
	10	8.2	7.6	7.2	3	2	1	2
	11	6.8	7.8	7.8	1	2	2	-
	12	7.6	6.0	7.2	3	1	2	3
	13	6.4	5.0	7.0	2	1	3	0
	14	5.8	5.8	7.2	1	1	2	-
	15	8.6	9.0	8.0	2	3	1	1
	16	6.6	7.0	6.4	2	3	1	1
	17	5.0	3.0	7.0	2	1	3	0
	18	6.0	4.2	7.0	2	1	3	0
	19	7.0	6.8	7.2	2	1	3	0
	20	5.8	3.8	6.4	2	1	3	0
CAP	1	3.6	3.6	5.4	2	2	1	-
	2	4.6	3.4	4.8	2	1	3	0
	3	5.0	3.0	6.4	2	1	3	0
	4	5.0	3.4	7.0	2	1	3	0
	5	1.4	2.0	3.0	1	2	3	1
	6	5.0	3.0	7.0	2	1	3	0
	7	4.4	3.0	4.8	2	1	3	0
	8	5.0	3.0	5.0	2	1	2	-
	9	4.4	3.2	4.6	2	1	3	0
	10	5.0	3.0	7.0	2	1	3	0
	11	5.0	3.0	6.8	2	1	3	0
	12	5.0	3.0	6.8	2	1	3	0
	13	3.0	2.6	1.6	3	2	1	2
	14	5.0	3.0	7.0	2	1	3	0
	15	4.0	3.0	5.0	2	1	3	0
	16	5.0	3.4	5.0	2	1	2	-
	17	3.4	2.8	4.0	2	1	3	0
	18	5.0	3.0	5.6	2	1	3	0
	19	6.4	5.0	6.6	2	1	3	0
	20	5.0	3.0	6.0	2	1	3	0
	21	5.0	3.0	7.0	2	1	3	0
	22	5.0	3.0	5.4	2	1	3	0

Table A.6: Relative frequencies of choices on the Pareto frontier sorted by categories

	Category	Decision j (Patient)														
		1 (1A)	2 (1B)	3 (1C)	4 (1D)	5 (1E)	6 (2A)	7 (2B)	8 (2C)	9 (2D)	10 (2E)	11 (3A)	12 (3B)	13 (3C)	14 (3D)	15 (3E)
FFS	PROMAX	1.00‡	0.25	0.00	0.15	0.32‡	0.72‡	0.16	0.10	0.10	0.25	0.21	0.00	0.00	0.00	0.06
FFS	PATMAX	1.00‡	0.25	0.32	0.20	0.16	0.28	0.11	0.25	0.15	0.40	0.42‡	0.83‡	0.56‡	0.67‡	0.24
FFS	SOCOPT	1.00‡	0.50	0.32	0.05	0.32‡	0.72‡	0.47	0.35	0.20	0.25	0.42‡	0.83‡	0.56‡	0.00‡	0.59
FFS	PAROTH	1.00‡	0.00	0.37	0.60	0.53	0.00	0.26	0.30	0.55	0.10	0.37	0.17	0.44	0.33	0.71
CAP	PROMAX	0.00	0.00	0.00	0.05	0.05	0.00	0.00	0.00	0.00	0.05	0.05	0.00	0.00	0.00	0.05
CAP	PATMAX	0.75‡	0.71‡	0.68‡	0.65‡	0.71‡	0.88‡	0.95‡	0.95‡	0.91‡	0.95‡	0.41	0.32	0.38	0.36	0.27
CAP	SOCOPT	0.75‡	0.71‡	0.68‡	0.65‡	0.71‡	0.88‡	0.95‡	0.95‡	0.91‡	0.95‡	0.18	0.18	0.10	0.23	0.23
CAP	PAROTH	0.25	0.29	0.32	0.30	0.24	0.13	0.05	0.05	0.09	0.00	0.36	0.50	0.52	0.41	0.45

‡ Note: Some patient benefit/profit pairs are covered by two (or more) *different* categories. In particular, the social optimal quantity (q_j^{soc}) coincides for some decisions with either the profit maximal quantity ($\hat{q}_j$) or the patient optimal quantity (q_j^*).
In FFS: $(\pi(\hat{q}_j), B(\hat{q}_j)) = (\pi(q_j^*), B(q_j^*)) = (\pi(q_j^{soc}), B(q_j^{soc}))$ for patient 1A ($j = 1$); $(\pi(\hat{q}_j), B(\hat{q}_j)) = (\pi(q_j^{soc}), B(q_j^{soc}))$ for patients 1E and 2A ($j = 5, 6$); $(\pi(q_j^*), B(q_j^*)) = (\pi(q_j^{soc}), B(q_j^{soc}))$ for for patients 3A,..,3D (j =11,..,14).
In CAP: $(\pi(q_j^*), B(q_j^*)) = (\pi(q_j^{soc}), B(q_j^{soc}))$ for for patients 1A,...,2E ($j = 1, \ldots, 10$).

Figure A.1: Pareto frontiers per decision in FFS

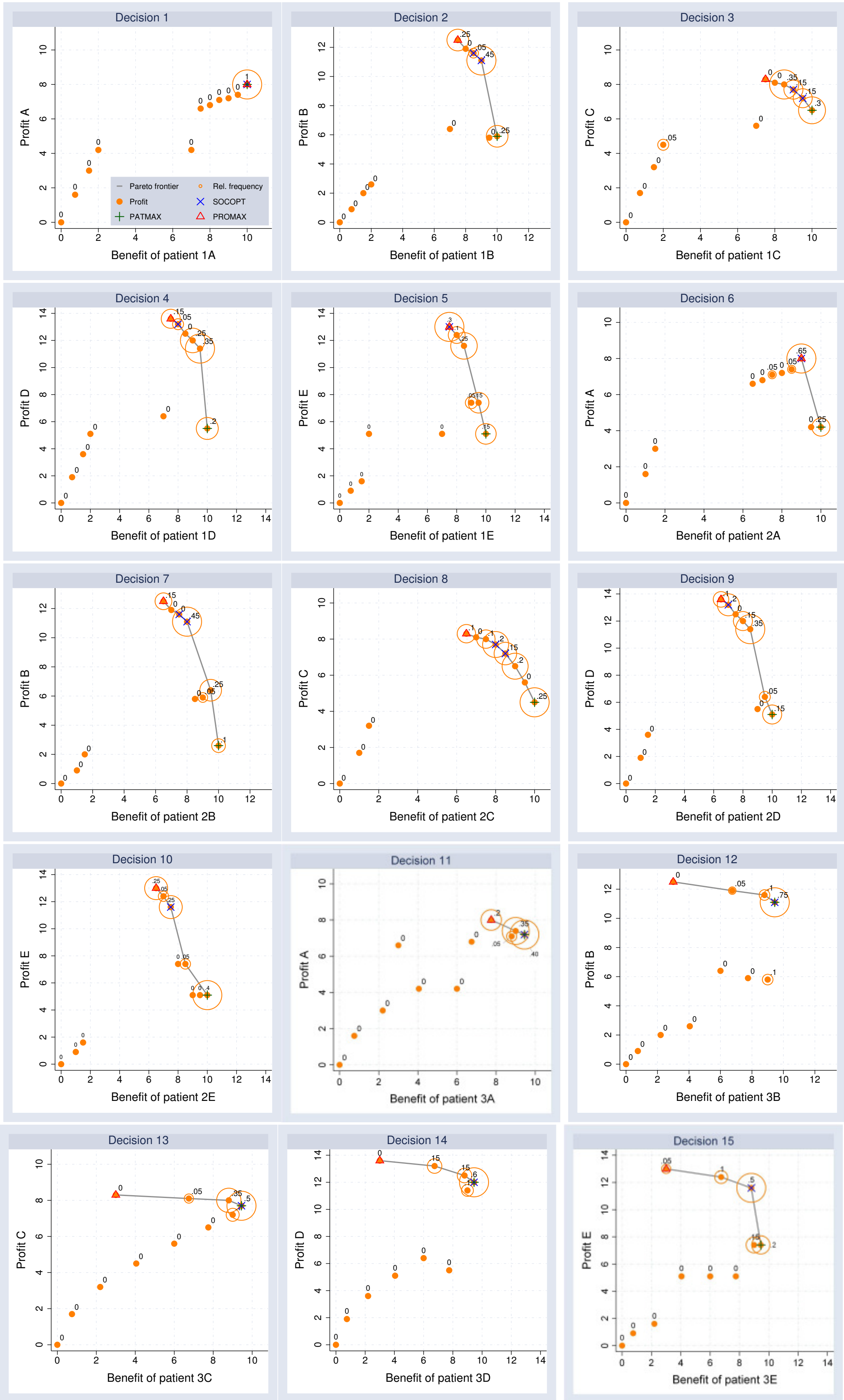


Figure A.2: Pareto frontiers per decision in CAP