Physicians' Provision Behavior under Different Payment Systems – An Experimental Investigation*

Heike Hennig-Schmidt, Reinhard Selten & Daniel Wiesen†

February 14, 2008

Abstract

A central concern in health economics is to understand influences of institutions on the behavior of actors on health care markets. In practice, effects from changing institutions, e.g. the payment system, in the course of a health care reform are *ex ante* not necessarily known to policy makers and may influence behavior in an undesired way. Main 'addressees' of reforms are health care providers, i.e. physicians, whose behavior is believed 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). Although empirical studies evidence that incentives from payment systems affect physicians' behavior, results are too contradictory for a definite conclusion about the direction of an effect to be drawn. Our study is meant to contribute to the research agenda put forward by Fuchs (2000), who suggests that using experimental economic methods may contribute beneficially to health economic research. We use a controlled laboratory experiment to improve the understanding of the institutional parameter 'payment system'. In our study, experimental physicians decide on the quantity of medical services under the two payment systems. Patients gain a monetary benefit from these services. No real patients participated in our experiment. To allow for other-regarding behavior of physicians the money corresponding to the benefits of all 'abstract' patients was donated to a charitable foundation caring for real patients. Our main finding is that patients are overserved under FFS and underserved under CAP. Financial incentives are not the only motivation for physicians' quantity decisions though. Patient benefit is of considerable importance as well. 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; controlled laboratory experiment; incentives
JEL-Classification: C91, I11

*We are grateful for helpful comments and suggestions by participants of the IZA Workshop on Behavioral and Organizational Economics and the summer school of the European Science Days in 2007. Financial support for Daniel Wiesen by the Konrad-Adenauer-Stiftung e.V. is gratefully acknowledged.

†Corresponding author: Daniel Wiesen, *B*onn*E*con*L*ab, Laboratory for Experimental Economics, Department of Economics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany; Phone: +49(0)228 739194, Fax: +49(0)228 739193, e-mail: daniel.wiesen@uni-bonn.de.

1 Introduction

While reforming a health care system, effects of variations of health care market institutions are *ex ante* not necessarily known to policy makers and may influence actors on the market in an undesired manner. As the general intention of health care reforms is to enhance efficiency, reduce costs and maintain or increase quality, a crucial role is attributed to health care providers, i.e. physicians. Their provision behavior is in turn believed to be influenced by incentives stemming from institutions like the payment system.

Many theoretical health-economic studies have highlighted the different incentives of commonly used payment systems. The two most prominent 'pure' payment systems analyzed are fee-for-service (henceforth FFS) and capitation (henceforth CAP). Under FFS, the physician is paid for each medical procedure or service dispended to a patient. FFS inherits an incentive to 'overserve' patients without considering costs (Newhouse 2002). When paid by CAP, physicians receive a fixed payment for each patient irrespective of the quantity of medical services provided. In contrast to FFS, CAP can reduce the utilization of health services, can lead to underprovision of medical services and to cream-skimming or even dumping of patients (Ma 1994).

There is empirical evidence that incentives from payment systems influence physicians' behavior, see e.g. Jennison & Ellis (1987), Stearns et al. (1992), Krasnik et al. (1990). Hutchinson et al. (1996) do not find differences, however. Thus, these results are not clear cut and remain too contradictory to draw a definite conclusion about the direction of an effect (Scott & Hall 1995, Gosden et al. 2001). In his article on the future of health economics, Fuchs (2000) makes the point that health economic research may largely benefit from incorporating methods of experimental economics. In fact, its use in health economics is negligible up to now.[1]

The purpose of our study is to contribute to the research agenda suggested by Fuchs (2000). We use a controlled laboratory experiment to improve the understanding of the institutional parameter 'payment system' likely to influence physicians' behavior. Our main focus is on *how* FFS and CAP influence a physician's provision of medical services abstracting from factors other than the payment system.

In our study, experimental physicians – all being 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. No real patients participated in our experiment. To allow for other-regarding behavior of physicians the money corresponding to the benefits of all abstract patients were donated to a charitable foundation caring for real patients. To the best of our knowledge our investigation is the first one tackling these issues by using experimental economic methods.

Our main finding is that physicians *are* influenced by the payment system in choosing more medical services when paid by FFS. Patients are overserved under FFS and underserved under CAP. Financial incentives are not the only motivation for physicians' quantity

[1]The only controlled laboratory experiment we know of that studies a health economic topic was conducted by Fan et al. (1998). The authors did not include an incentive to care for the patient, however.

decisions; patient benefit is of considerable importance as well. Given our experimental setup, 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.

Our paper is organized as follows. Section 2 introduces a simple theoretical model of physician behavior and the model's implications for the two payment systems are derived (Subsection 2.1). This model is used because it includes components which can also be found in the present experimental design. Moreover, this section reviews the empirical literature (Subsection 2.2), discusses the need for a labarotory experiment and states our research questions (Subsection 2.3). Section 3 the experimental design as well as the experimental procedure are described. Section 4 provides statistical analysis of subjects' behavior and compares experimental results across payment systems. Section 5 concludes.

2 Physicians' incentives from payment systems

The impact of the payment system on physicians' provision behavior has been the subject of various theoretical and empirical investigations. This section provides a brief overview on the main findings. Before reviewing the literature, we use the seminal model by Ellis & McGuire (1986) as a workhorse to analyze the incentives from FFS and CAP.

So far, no consensus has been reached on how to formally model physician agency (Choné & Ma 2006). The lack of consensus might originate from the fact that the fundamentals of the problem, e.g. motives like medical ethics, benevolence towards the patient and power, and imperfect information are very complex. Institutions like payment system or health insurance and their complexity overlaying phyisicans' decisions add to the challenge. A simple profit-maximizing approach obviously does not capture the full extent of this complexity. Conventional modeling of the physician-patient interaction relies on profit maximization, however (e.g. McGuire & Pauly 1991). Frequently, a complete information framework is added. McGuire (2000) critically notes that even though there is no agreeable alternative to model the physician in the conventional way, this approach is not well-accepted in health economics.

In the recent literature, several authors depart from modeling physicians as pure profit maximizers by allowing for patient benevolence in the physician's utility function, see e.g. Ellis & McGuire (1986, 1990), Chalkley & Malcomson (1998), Ma (2004), Jack (2005) and Choné & Ma (2006).

2.1 A basic model of physician behavior

In Ellis & McGuire's basic model of 1986, the physician[2] decides on the quantity of medical services as an agent of the patient and the hospital. Following Newhouse (2002), we apply their model to a primary care physician. The physician is assumed to be concerned about her own profit π and patient benefit B the latter depending on the quantity of medical

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

services $q \in [0, Q]$. 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).[3] Physician's effort positively entering the patient benefit function is constant.

The utility of the physician is as follows

$$U(\pi, B) = \pi + \alpha \cdot B(q), \tag{1}$$

with $\alpha \in [0, 1]$. α can be interpreted as an index for the physician's benevolence, i.e. the degree she acts on the patient's behalf when deciding on q. $\alpha = 0$ means, the physician does not act on the patient's behalf. If $\alpha = 1$, the physician equally weighs own profit and patient benefit being benevolent towards the patient. For $0 \leq \alpha < 1$, the physician acts partially benevolent towards the patient.

The benefit function is assumed to be strictly concave on the interval $[0, Q]$, with a global optimum $B'(q^*) = 0$, $B'(q) > 0$ for $q \in [0, q^*)$, $B'(q) < 0$ for $q \in (q^*, Q]$ and $B''(q) < 0$, see also Ellis & McGuire (1990), Ma (2004) and Choné & Ma (2006). The negative second order derivative can be interpreted as stemming from the following factors. Patients may have a declining marginal valuation of health care because the marginal benefit is lower as more health care is consumed.[4] Moreover, when a higher health status is gained the marginal utility of health status itself falls.

The patient is assumed to be passive and fully insured accepting each quantity of medical services provided. Thus, the physician's quantity decision is not restricted by a patient's demand for medical services.[5]

Physician's profit is determined by the payment (R) the physician receives minus his cost per treatment ($C(q)$) yielding $\pi = R - C(q)$. The cost function $C(q)$ is strictly increasing and convex. The efficient quantity of medical service (q^{**}) is chosen such that $\alpha B'(q) = C'(q)$.

The remuneration R may consist of a capitation component A and a fee-for-service component $p \cdot q$, where p is a vector of prices the physician receives according to the vector of services q provided. Thus, physician's profit can be written as

$$\pi = A + p \cdot q - C(q). \tag{2}$$

Plugging (2) into (1) gives

$$U = A + p \cdot q - C(q) + \alpha \cdot B(q). \tag{3}$$

[3] Following this argumentation, Arrow (1963) emphasizes the importance of professional ethics limiting payment incentives. Also, McGuire (2000) and Newhouse (2002) consider professional ethics a motive for physician behavior.

[4] Think of the decreasing value of further X-ray examinations if the first one has already identified pneumonia.

[5] Assuming a passive patient is crucial because it eliminates demand-side effects as e.g. patients' co-payments likely to influence the physician's quantity decision. Additionally, collusion between patient and physician are avoided as modeled in Ma & McGuire (1997).

Incentives from FFS

In traditional health care systems, medical care delivery is priced on a fee-for-service basis. We abstract from supply-side cost sharing[6] but assume complete cost reimbursement with the fees being equal or above marginal cost, $p \geq C'(q)$. The physician's remuneration equals $R = p \cdot q$ and her utility is

$$U = p \cdot q - C(q) + \alpha \cdot B(q). \tag{4}$$

For the following analysis of physician behavior we assume $p > C'(q)$. Simple differentiation shows that the efficiency condition does not hold in a fee-for-service payment system.[7]

Predictions on the quantity provided vary according to the physician's degree of benevolence towards the patient.

- A purely profit-maximizing physician ($\alpha = 0$) chooses q independent of the patient's interest. Thus, she chooses the maximum available amount of services $q^{FFS}_{(\alpha=0)} = Q$.
- For $\alpha = 1$, the physician chooses $q^{FFS}_{(\alpha=1)}$ such that $B'(q) = C'(q) - p$. According to the first order condition, $q^{FFS}_{(\alpha=1)} > q^*$.[8] Hence, even a benevolent physician chooses a quantity of medical services larger than optimal for the patient.
- For $0 < \alpha < 1$, the physician gives less value to the patient's benefit than to her own profit and chooses $q^{FFS}_{(0<\alpha<1)}$ such that $B'(q) = (C'(q) - p)/\alpha$. Thus, this physician provides medical services according to $q^{FFS}_{(\alpha=1)} < q^{FFS}_{(0<\alpha<1)} < q^{FFS}_{(\alpha=0)}$.

Our analysis shows that regardless of the physician's degree of benevolence, FFS leads to oversupply of medical services in that $q^{FFS} > q^* > q^{**}$. The quantity provided is larger than the patient's optimal and the efficient quantity.

Incentives from CAP

Another form of physicians' remuneration is a prospective capitation payment. In a given period, the physician is paid a lump sum for each registered patient independent of the quantity of medical services she provides.

The physician's utility function under CAP is

$$U = A - C(q) + \alpha \cdot B(q) \tag{5}$$

with $A > C'(q)$. For a given value of α, the efficient quantity of medical services q^{**} is such that the first-order condition $\partial U/\partial q = 0 \Leftrightarrow C'(q) = \alpha B'(q)$ holds.

- The incentive inherent in a capitation payment leads the profit-maximizing physician ($\alpha = 0$) to increase the difference between the lump-sum payment and the personal

[6] Supply-side cost sharing means that the fees per unit of treatment the third-party payer pays to the provider are lower than the cost per unit of treatment, $p < C'(q)$.

[7] If fees are set at marginal cost and $C'(q) = c$, physician's profit equals $\pi = (p - c) \cdot q \equiv 0$. If $\partial\pi/\partial q = 0$, the physician treats the fully insured patient according to $B'(q) = C'(q) - p = 0$. Thus, in line with Newhouse (2002), the physician will choose q such that any expected positive benefit for the patient is provided, irrespective of cost.

[8] From $B'(q) = C'(q) - p$ and $p > C'(q)$ follows $B'(q) < 0$, i.e. $q > q^*$.

costs per medical service ($\text{argmax}_q\{A - C(q)\}$). The physician sets $q^{CAP}_{(\alpha=0)} = 0$, i.e. no medical service will be delivered to the patient.

- A benevolent physician ($\alpha = 1$), chooses $q^{CAP}_{(\alpha=1)}$ such that $B'(q) = C'(q)$; i.e. she chooses the efficient quantity $q^{CAP}_{(\alpha=1)} = q^{**}$.
- For $0 < \alpha < 1$, the physician decides on $q^{CAP}_{(0<\alpha<1)}$ according to $B'(q) = C'(q)/\alpha$.

Physicians provide medical services according to $q^{CAP}_{(\alpha=0)} < q^{CAP}_{(0<\alpha<1)} < q^{CAP}_{(\alpha=1)} = q^{**}$. As $q^* > q^{**}$, all physicians underserve patients under CAP.

The most important result of the above analysis is that regardless of the physician's degree of benevolence towards the patient, the quantity she provides under FFS is larger than under CAP, i.e. $q^{FFS} > q^{CAP}$.

2.2 Empirical literature

In this section we briefly summarize the relevant empirical literature on physician behavior under different payment systems, in particular under FFS and CAP. Jennison & Ellis (1987) using data from the US find the same physicians to provide more visits under a generous FFS system than under CAP. A similar result is reported by Stearns et al. (1992) who find a reduction in hospitalizations but increases in length of hospital stay and number of ambulatory visits. They conjecture that increases may be due to CAP payment for primary care physicians and a reduced FFS schedule for specialists leading to a greater number of referrals. On the contrary, Hutchinson et al. (1996) do not find differences when comparing hospital utilization rates in Ontario (Canada) under FFS and CAP.

Krasnik et al. (1990) analyze behavior of general practitioners in Denmark when the system is varied from a (pure) lump-sum payment to capitation supplemented by a fee-per-item component. They find diagnostic and curative services to increase and the number of referrals to secondary care and hospitals to decrease. Krasnik et al. analyzed a random sample taken from the participating physicians. Their study, however, may be biased by self-selection of practitioners.

In a randomized controlled study, Davidson et al. (1992) investigate behavior 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. Physicians paid by capitation were responsible for almost all cost of the children enrolled with them (fundholding). The authors find the frequency of primary care visits in the high FFS-group to be higher than in the capitation group. This seems to provide evidence that CAP-physicians constrain the quantity of medical services in order to reduce their costs. The fundholding regulation 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.

Iversen & Lurås (2000) analyze referral rates from primary to secondary care revealed by Norwegian general practitioners when the payment system was changed from a practice allowance component[9] complemented by a FFS-payment to a CAP-system with a lower

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

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 attributed to CAP but rather to the lower FFS-component.

In a more recent empirical study, Dumont et al. (2007) analyze 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. Physicians could voluntarily choose one of the two systems. Their results suggest that physicians did react to payment incentives by reducing the volume of (billable) services under the mixed remuneration system. Moreover, physicians switching to the mixed system increased the time spent per service and per non-clinical services such administrative and teaching tasks (services that are important to insure the quality of health care but not remunerated under FFS). This suggests a quantity-quality substitution when physicians care for patients.

2.3 Research questions

Our main research goal is to improve the understanding on how the institutional parameter 'payment system' influences physicians' behavior. To this end, we make use of experimental economics methods by running a controlled laboratory experiment. Why do we use experimental economics to pursue our research goals? Experimental economics is a valid research method because of a variety of advantages compared to field data and questionnaire studies (see Falk & Fehr 2003, Davis & Holt 1993).

Experimental data is created for scientific purposes under controlled conditions. It is gathered in experimental sessions in which human subjects make *real* decisions in *economically relevant* decision situations supplied with *monetary incentives*. The reason for paying participants is that subjects in behavioral decision making are likely to behave differently when monetary consequences are involved compared to hypothetical situations (Hertwig & Ortmann 2001, Camerer 2003).

Experimental conditions can be varied in a controlled manner. Exogenous *ceteris paribus* variations (of the payment system) can be easily implemented, variables of interest (physicians' behavior) can be controlled. Therefore, changes in behavior can be attributed to these modifications. Finally, different experimenters can repeat the same experiment under comparable conditions in order 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 that the researcher usually cannot control. Based on their meta-study, Gosden et al. (2001) stress that field studies face the difficulty of multiple and unobservable influences on physicians' behavior.[10] They tend to be context-specific limiting the application of results to other settings or rendering a

[10] These are among others institutional parameters, physicians' characteristics, uncertainty about the impact of medical services provided (Arrow 1963), fear of malpractice litigation (Kessler & McClellan 1996) as well as patient characteristics like health status (see the literature on cream-skimming, e.g. Newhouse 1996, Barros 2003) or type of insurance (Eisenberg 1986). Therefore, constant patient populations during a transition of payment systems is important for the validity of results (Hutchinson et al. 1996) but can most often not be guaranteed.

generalization of results difficult. The authors also point out that several field studies suffer from methodological problems. Most importantly, in some studies more than one component of the payment system are varied simultaneously making causal inferences difficult or even impossible. Nonetheless, Gosden et al. admit some empirical evidence that the payment system affects physician behavior.

Despite the advantages of experimental economics, objections like non-representative student subject pools, low incentives, the small number of participants and the simplistic environment should be taken seriously. Yet, careful experimentation can circumvent many of these objections (Falk & Fehr 2003). We are aware that our experiment is extremely simplistic as we abstract from factors other than the payment system. In reality, a physician's decision situation is definitely more complex. Yet, as the goal of the present study is to highlight fundamental consequences of the payment system for physicians' behavior we think simplicity to be an advantage. Laboratory experimentation should be regarded as complementary to theoretical analysis and other methods of empirical investigation. Its contribution may help to draw a more precise picture of physicians' provision behavior.

The main focus of our study is on how the pure payment systems FFS and CAP influence an experimental physician's provision of medical services. Recall that experimental physicians decide on the quantity of medical services. We look at provision behavior from the physician's and from the patient's perspective.

Our first research question is concerned with behavior in FFS. Given our experimental parameters, do experimental physicans tend to behave according to what theory quantitatively predicts (subsection 2.1)? Do they choose a quantity of medical services q^{FFS} larger than the patient's optimal quantity q^* if the profit-optimal quantity $\hat{q}$ exceeds q^*? Taking q^* as the benchmark for the right (best) medical treatment, we expect patients to be overserved under FFS.

Our second research question deals with behavior under CAP. According to theoretical predictions we conjecture that physicans choose a quantity of medical services q^{CAP} lower than the patient's optimal quantity q^*. Taking q^* again as a measure for the best medical treatment, we assume patients to be underserved under CAP.

In both payment systems, we are interested in the interplay between the mode of payment and patients' health status. Do patient types benefit from FFS and CAP in the same way?

Our third research question is concerned with a comparison of behavior under FFS and CAP. We expect experimental physicans in FFS to choose more medical services than subjects in CAP do. Such behavior would be in line with the empirical findings of Krasnik et al. (1990) or Dumont et al. (2007).

Our forth research question deals with physician's profit and patient benefit. We are interested whether besides their own profit, experimental physicians care for their patients and take the patient benefit into account when making their quantity decisions. Given the professional code of medical ethics physicians are obliged to, we expect our experimental physicians not to behave in a completely self-interested manner. Yet, also their own profit should be of considerable importance.

Our last research question involves analysing the tradeoffs between own profit and patient benefit the experimental physicians are faced with. According to the experimental parameters, 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. Does behavior with regard to tradeoffs vary in the two payment systems? Do subjects differ in their choices with regard to tradeoff? Can a classification of behavior help us to understand differences in decision making like e.g. in Selten et al. (1997) and Fischbacher et al. (2001).

3 Experimental design and procedure

3.1 Design and parameters

The focus of our study is on physicians' provision behavior under two different payment systems. No other experimental parameter than the payment condition is varied. We chose an experimental design allowing for a controlled and isolated analysis, i.e a *ceteris paribus* variation, and a between-subject comparison.

Subjects participating in our experiment are exclusively medical students likely to become physicians in the future. This is important in the context of our experiment as each subject is allocated to a physician's role deciding on the quantity of medical services to be provided for given patient. We call our experimental subjects physicians according to the role they play in our experiment. The role of patients will be explained shortly.

The experiment consists of two treatments, FFS and CAP (see Table 1). In each treatment, physicians are remunerated for their provision of medical services in a different way. In FFS, physicians are paid by fee-for-service, i.e. they receive a fee for each unit of medical service provided. In CAP, physicians are paid a lump-sum payment (capitation) per patient independent of the number of medical services they provide.

Table 1: Experimental treatments

Treat.	Payment condition	Number of sessions	Number of independent observations
FFS	Fee-for-service	1	20
CAP	Capitation	2	22

In both treatments, physicians decide on the quantity of medical services $q \in \{0, 1, \ldots, 10\}$ for five given abstract illnesses $h = \mathrm{A}, \mathrm{B}, \ldots, \mathrm{E}$ of three different patient types $k = 1, 2, 3$. Each combination of patient type and illness represents a specific patient $kh = 1\mathrm{A}, 1\mathrm{B}, 1\mathrm{C}, \ldots, 2\mathrm{E}, 3\mathrm{E}$ (see Table 2). Each decision $j = 1, \ldots, 15$ simultaneously determines the physician's own profit and the benefit of a given patient. We will elaborate on these experimental variables in the following.

The range of services physicians can choose from may be interpreted as those eligible for a patient contracting with a certain health plan. We did not characterize illnesses in

Table 2: Order of decisions

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

real terms because this turned out not to be feasible. The patient is assumed to be passive and fully insured accepting each medical service chosen by a physician. All experimental parameters except the quantity of services are measured in Taler, our experimental currency, given an exchange rate of 1 Taler = 0.05€.

Physicians' remuneration

In FFS, physicians receive a fee for each unit of medical service provided. Fees differ across services. Remuneration $R_h(q)$ increases in the quantity of medical services chosen (Table 3).[11]

In CAP, physicians are paid a lump-sum payment R per patient independent of their

Table 3: Physicians' remuneration $R(q)$

		Quantity (q)										
		0	1	2	3	4	5	6	7	8	9	10
FFS	$R_A(q)$	0.00	1.70	3.40	5.10	5.80‡	10.50	11.00	12.10	13.50	14.90	16.60
FFS	$R_B(q)$	0.00	1.00	2.40	3.50	8.00	8.40	9.40	16.00	18.00	20.00	22.50
FFS	$R_C(q)$	0.00	1.80	3.60	5.40	7.20	9.00	10.80	12.60	14.40	16.20	18.30
FFS	$R_D(q)$	0.00	2.00	4.00	6.00	8.00	8.20	15.00	16.90	18.90	21.30	23.60
FFS	$R_E(q)$	0.00	1.00	2.00	6.00	6.70	7.60	11.00	12.30	18.00	20.50	23.00
CAP	R	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 $R_A(q)$ 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.*

quantity decision. R is set to 12 Taler slightly above the average maximum profit per patient in FFS (11.06).

Patient benefit

Patients gain a benefit from medical services, the patient benefit B. In our study, the patient benefit is measured in monetary terms. Note that no real patients participated in our experiment. Yet even with abstract patients, we wanted to allow for a motivation of other-regarding behavior a physician may experience while treating a real patient. To this end, the benefits of all abstract patients aggregated over all decisions of all physicians were donated to a charitable foundation caring for real patients – the *Christoffel Blindenmission*. This foundation is engaged in treating ophthalmic patients mainly in developing countries.

[11] Different fees for different kinds of services can be found in practice e.g. in Germany. The German GOÄ (*Gebührenordnung für Ärzte*) lists medical services and the respective fees.

To gain credibility that the donation was actually transferred to the charity organization, a monitor was randomly selected from the participating subjects and carried out the donation transfer procedure in each session like in Eckel & Grossman (1996). A copy of the instructions is included in Appendix A.1.

Patient benefits vary across patient types. This reflects the heterogenity of the patient

Table 4: Patient benefit $B_k(q)$

	Quantity (q)										
	0	1	2	3	4	5	6	7	8	9	10
$B_1(q)$	0.00	0.75	1.50	2.00	7.00	10.00‡	9.50	9.00	8.50	8.00	7.50
$B_2(q)$	0.00	1.00	1.50	10.00‡	9.50	9.00	8.50	8.00	7.50	7.00	6.50
$B_3(q)$	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^* providing the patient with the highest benefit $B_k(q_j^*)$ from medical services.*

population treated by a physician in reality e.g. with regard to patients' states of health or different severities of illness. Table 4 shows patient benefits $B_k(q)$ according to the quantity of medical services provided. A common characteristic of $B_k(q)$ is a global optimum $q^* \in [0, 10]$. 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 ($j = 11, \ldots, 15$).[12] After having reached the optimum, $B_k(q)$ declines because providing too many medical services contributes negatively to a patient's benefit at the margin. As there is a unique optimal q_j^* for each decision j (patient kh), overprovision or underprovision can be identified.

Pysicians' profit

Further parameters relevant for physicians' decisions are profit (costs). Like real doctors, the experimental physicians have to bear costs depending on the quantity of medical services they choose. The costs are kept constant across treatments and follow the convex function $c(q) = 0.1 \cdot q^2$ (see Table 5).[13]. In FFS, profit varies across illnesses because

Table 5: Physicians' costs $c(q)$

	Quantity (q)										
	0	1	2	3	4	5	6	7	8	9	10
$c(q)$	0.00	0.10	0.40	0.90	1.60	2.50	3.60	4.90	6.40	8.10	10.00

fees differ for patients and cost parameters are kept constant; in CAP, however, profit is constant across illnesses and patient types (see Table 6).

For all decisions j of FFS, except for $j = 1$ (patient 1A), experimental parameters imply the patient's optimal quantity of medical services q_j^* to differ from the quantity $\hat{q}_j$ providing the maximal profit to the physician. For $j = 1$, $\hat{q}_j$ coincides with q_j^* at $q_1 = 5$.

[12]Patient type 2 (3) can be considered as needing a relatively low (high) quantity of medical services to gain her health optimum whereas patient type 1's optimum is in between.

[13]A convex cost function is assumed in several theoretical papers (Ma 1994, Ma 2004, Choné & Ma 2006) as well as in Fan et al. (1998)

For $j = 11$ (patient A3), $5 = \hat{q}_j < q_j^* = 7$.

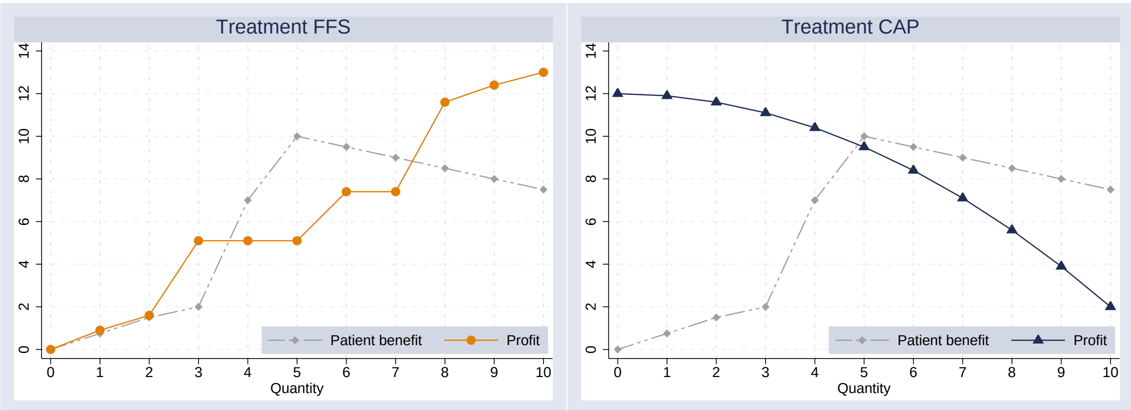
In decisions $j = 2, \ldots, 15$, the physician encounters a tradeoff between patient's optimum and own profit maximization. She foregoes own profit when increasing the patient's benefit and vice versa. Physicians face decisions where choosing more medical services implies a large increase in patient benefit but only a marginal decrease in own profit like e.g. in decision $j = 2$ (patient 1B) in FFS. Choosing $q = 5$ instead of 4 provides patient 1B with a benefit of 10.00 instead of 7.00 (Table 4) while the physician's profit decreases from 6.40 to 5.90 Taler only (Table 6). Note that a higher q does not neccessarily imply a higher profit. In decisions $j = 1, 6, 11$ (patients 1A, 2A, 3A) a lower level of services provides a higher profit. In CAP, $\hat{q}_j = 0$ for each decision ($j = 1, \ldots, 15$). Higher or maximal patient benefits can only be achieved by physicians' foregoing own profit. An illustration provides Figure 1 for patient 1E (decision $j = 5$).

Table 6: Physicians' profit $\pi(q)$

		Quantity (q)										
		0	1	2	3	4	5	6	7	8	9	10
FFS	$\pi_A(q)$	0.00	1.60	3.00	4.20	4.20	8.00‡	7.40	7.20	7.10	6.80	6.60
	$\pi_B(q)$	0.00	0.90	2.00	2.60	6.40	5.90	5.80	11.10	11.60	11.90	12.50‡
	$\pi_C(q)$	0.00	1.70	3.20	4.50	5.60	6.50	7.20	7.70	8.00	8.10	8.30‡
	$\pi_D(q)$	0.00	1.90	3.60	5.10	6.40	5.50	11.40	12.00	12.50	13.20	13.60‡
	$\pi_E(q)$	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

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

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



3.2 Procedure

The computerized experiment was conducted in *B*onn*E*con*L*ab, the Laboratory for Experimental Economics at the University of Bonn. We used the software z-Tree (Fischbacher 2007). 42 medical students participated in the two treatments – 20 in FFS and 22 in CAP. Subjects made their decisions anonymously at their computer screens without any communication.

Upon arrival, each subject randomly drew a number indicating his/her cubicle where he/she remained seated during the whole experiment. First, the experimenter read the instructions aloud.[14] Then, subjects were given time for clarifying questions which were asked and answered in private. In order to check for subjects' understanding of the experiment, they had to answer three test questions structured like the actual experiment but with different parameter values. Subjects needed three different quantities of medical services to answer the questions. To avoid any priming by pre-selected quantities the three numbers q were randomly drawn from the interval $[0, 10]$ from a box and announced by the experimenter. The experiment was not started unless all participants had answered all test questions correctly.

In both treatments, each participant was assigned the role of a physician having to make 15 decisions ($j = 1, \ldots, 15$) on the quantity of medical services. The sequence of decisions (patients) was predetermined and kept across treatments (see Table 2). 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's role was assigned to one of the participants by random draw. After the experiment, subjects were paid in private according to their performance.

Similar to the procedure in Eckel & Grossman (1996), the monitor had to verify, by a signed statement available to all participants, that a check for the total amount corresponding to the aggregated patient benefits was written and sealed in an envelope addressed to the charity. The monitor and experimenter then walked together to the nearest mailbox and deposited the envelope.

The experiment lasted for about 45 minutes. On average subjects earned 6.88€ in FFS and 7.42€ in CAP. In total, 273.68€ were transferred to the *Christoffel Blindenmission*, 6.62€ per participant in FFS and 6.42€ per participant in CAP.

4 Results

4.1 Physicians' provision behavior

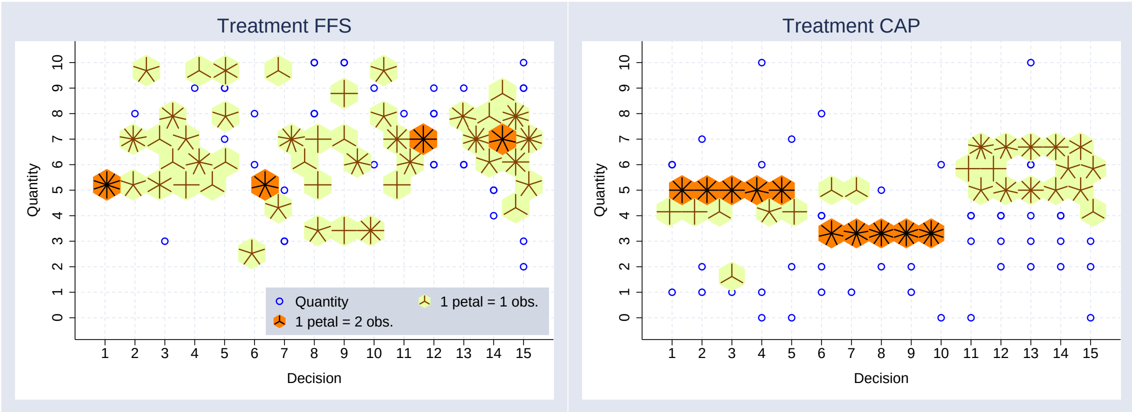
In this section, we give a detailed analysis of physicians' behavior, both from the physician's and patient's perspective. In particular, we analyze physician i's quantity decisions q_{ij} for FFS and CAP separately. To get a first glimpse of behavior in both treatments see Figure 2. The same will be done for i's deviations from the patient optimal quantity ($q_{ij} - q_j^* \equiv \mu_i$). Considering q_j^* the benchmark for providing the ideal quantity of medical services for a patient, each $q_{ij} > q_j^*$ ($q_{ij} < q_j^*$) indicates overprovision (underprovision) yielding a lower benefit for the patient. Recall that $q_j^* = 5$ for $j \in [1, 5]$, $q_j^* = 3$ for $j \in [6, 10]$ and $q_j^* = 7$ for $j \in [11, 15]$.

4.1.1 Behavior in FFS

Our first research goal is concerned with behavior under FFS. Will patients on average be overserved as theory predicts given our experimental parameters? To answer this question

[14] For detailed instructions see Appendix A.1.

Figure 2: Absolute frequencies of quantity decisions per patient



we analyze the quantity of medical services provided for each patient kh. In addition, we analyze the impact of the payment system on patients' health status with regard to patient types. Remember that for $j = 1$ (patient 1A), $\hat{q}_j = q_j^*$, and for $j = 11$ (patient 3A), $\hat{q}_j < q_j^*$.

Averaged over all physicians and all patients, a mean quantity of medical services $\overline{q}^{FFS} = 6.60$ (median $\tilde{q}^{FFS} = 7.00$) is chosen (see Table 7). Figure 3 shows average quantities for each decision (patient) separately.

Table 7: Quantity decisions q in FFS and CAP

	Mean ($\overline{q}$)	Median ($\tilde{q}$)	SD	Total number of decisions
FFS	6.60	7.00	1.85	300
CAP	4.40	5.00	1.64	330

We first take a closer look at how patients are treated. To this end, we analyze the quantity of medical services provided for each patient kh (decision j) averaged over all physicians ($\overline{q}_j$).

Result 1. *In FFS, patients are overserved compared to their optimal treatment.*

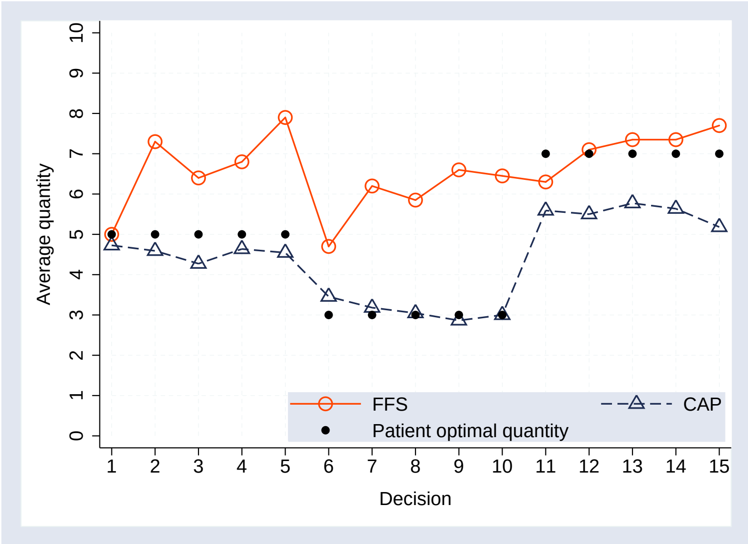
SUPPORT: Figure 3 shows average quantities $\overline{q}_j$ ($\overline{q}_j = \sum_{i=1}^{20} q_{ij}/20$) to be larger than q_j^* for the 13 patients with $\hat{q}_j > q_j^*$ (patients 1B, ..., 2E, 3B, ..., 3E). For patient 1A ($j = 1$), all physicians i chose $q_{i1} = q_1^* = 5$, whereas for patient 3A ($j = 11$), $\overline{q}_{11} < q_{11}^*$. Testing over all patients, we find $\overline{q}_j$ to be highly significantly larger than the patient optimal quantity q_j^* ($p = 0.0021$, Wilcoxon signed ranks test, two-sided).

The second result is concerned with decisions of the individual physician. We analyze the (averaged) quantity of medical services each physician i provides for the 15 different patients.

Result 2. *Physicians in FFS provide quantities of medical services larger than q_j^*.*

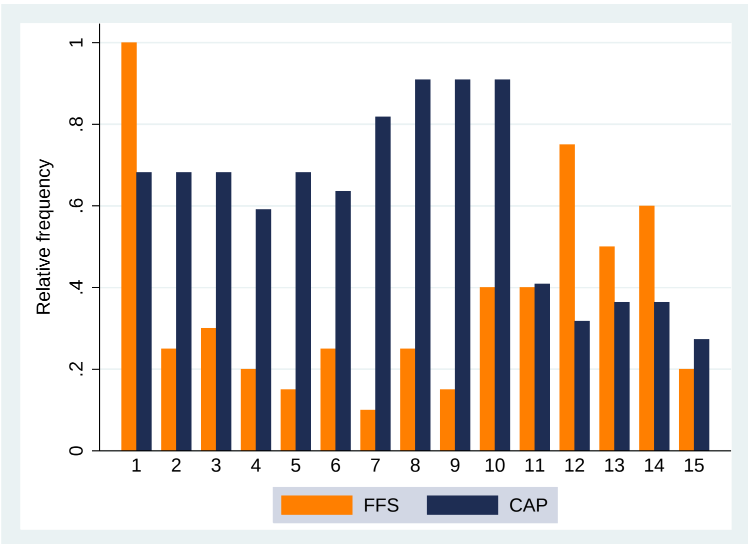
SUPPORT: Table A.1 in Appendix A.2 shows physicians' mean quantity decisions ($\overline{q}_i$) and

Figure 3: Average quantity of medical services per decision (patient)



the mean deviations from the patient optimal quantity q_j^*, $\overline{\mu}_i = \sum_{j=1}^{15}(q_{ij} - q_j^*)/15$. For 17 out of the 20 physicians, $\overline{\mu}_i$ is positive and zero for the remainder. Thus, physicians overserve in FFS in that highly significantly more physicans provide patients with medical services *on average* larger than q_j^* ($p = 0.003$, binomial test, two-sided). Even stronger support is provided by test statistics of a two-sided Fisher-Pitman permutation test for paired samples for individual decisions. For 16 of the 20 physicians, the null hypothesis of $q_{ij} = q_j^*$, $\forall j \in [1, 15]$ can be rejected. These physicians chose quantities significanly larger than q_j^* (see Table A.2). Thus, highly significantly more physicans provide patients with medical services that are *significantly* larger than q_j^* ($p = 0.012$, binomial test, two-sided).

Figure 4: Relative frequencies of patient optimal quantity choices



Next we investigate the impact of patients' characteristics on physicians' behavior. In particular, we analyze whether certain types of patients are overserved. Thereunto, we compute for each patient the number of physicians chosing the patient's optimal quantity (q_j^*) and those deviating from it ($\neg q_j^*$); see Table A.3. A graphical illustration of relative frequencies of q_j^*-choices provides Figure 4.

Result 3. *Overprovision in FFS depends on patient types.*

SUPPORT: Except for patient 1A (decision 1) where $q_{ij} = q_j^* = \hat{q}_j$, all patients of type 1 and 2 are overserved in that the number of physicians choosing $q_{ij} > q_j^*$ is larger than the number of physicians choosing $q_{ij} \leq q_j^*$ (see Figure 2). This is significant for 3 (4) patients of type 1 (2) (binomial test, two-sided; see line I/FFS in Table A.4). 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.[15] Note that only patients of type 3 are underprovided, except for one decision of a single physician ($q_{i=11,j=3}$).

When comparing the average deviation $\overline{\nu}_j$ for each patient ($\overline{\nu}_j = \sum_{i=1}^{20}(q_{ij} - q_j^*)/20$), the above differences appear to be corroborated. $\overline{\nu}_j$ is not larger than 1.00 for patient type 3. It varies between 1.80 and 2.90 for patient type 1 and between 1.70 and 3.60 for patient type 2 (Table A.5).

Results 1 and 2 suggest that our experimental physicians in FFS behave like we expected them to do. Patients are overserved in that subjects on average choose quantities of medical services q^{FFS} larger than the patient's optimal quantity q^*. Physicians' choices are heavily dependent on patient types, however (Result 3). When the difference between q_j^* and $\hat{q}_j$ becomes smaller, the number of optimal choices increases.

4.1.2 Behavior in CAP

Our second research goal deals with behavior under CAP. We are interested in whether experimental physicans tend to underserve patients. We proceed like in FFS by analyzing the quantity of medical services provided for each patient kh. In addition, we investigate the impact of CAP on physicians' decisions with regard to patient types. Recall that $0 = \hat{q}_j < q_j^*$ for all decisions j (patients kh).

Averaged over all physicians and all patients, a mean quantity of medical services $\overline{q}^{CAP} = 4.40$ (median $\tilde{q}^{CAP} = 5.00$) is chosen (see Table 7). Figure 3 shows average quantities for each decision (patient) separately. We first investigate how patients are treated.

Result 4. *In CAP, patients are underserved compared to their optimal treatment.*

SUPPORT: Figure 3 shows average quantities $\overline{q}_j$ to be smaller than q_j^* for 11 of the 15 patients. Patients 2A, 2B and 2C are overserved whereas only patient 2E is optimally

[15] On average, 15.25 (15.40) physicians overserve patients of type 1 (2), but only 6.20 overserve those of type 3. An average number of 9.8 physicians optimally treat patients of type 3, but only 4.5 (4.6) choose q_j^* for patients of type 1 (2).

treated on average. Testing over all patients kh, we find $\overline{q}_j$ to be significantly smaller than q^*_j (p=0.0105, Wilcoxon signed ranks test, two-sided).

As in FFS the next result is concerned with decisions of the individual physician. We analyze the quantity of medical services each of the 22 experimental physicians provides averaged over the 15 patients ($\overline{q}_i$).

Result 5. *Physicians in CAP provide quantities of medical services below* q^*_j.

SUPPORT: Table A.1 provides support for Result 5. $\overline{\mu}_i$ is negative for 16 physicians. Among the remaining 6 physicians $\overline{\mu}_i > 0$ for $i = 4, 19$ and $\overline{\mu}_i = 0$ for $i = 6, 10, 14, 21$. Thus, physicians underserve in CAP in that weakly significantly more physicians provide patients with medical services *on average* smaller than q^*_j ($p = 0.0525$, binomial test, two-sided). Test statistics of a two-sided Fisher-Pitman permutation test for each individual physician's decisions indicate a similar picture. Table A.2 shows that for 12 physicians the null hypothesis of $q_{ij} = q^*_j$, $\forall j \in [1, 15]$, can be rejected at a 10 percent level. These physicians chose quantities (weakly) significanly lower than q^*_j.

Next we investigate whether underprovision depends on patient types. As in FFS, we compute the number of physicians chosing the patient's optimal quantitiy q^*_j, and we calculate the number of physicians choosing $\neg q^*_j$ (see Table A.3). See also Figure 4 for relative frequencies of q^*_j.

Result 6. *Underprovision in CAP depends on patient types.*

SUPPORT: All patients of type 1 and 2 are treated in a rather benevolent manner in that the number of physicians chosing q^*_j is larger than the number of physicians chosing $\neg q^*_j$ (Figure 2). This is significant for 4 patients of type 2 (binomial test two-sided; see line I/CAP in Table A.4).[16] Patients of type 3 are underserved; the number of physicians chosing $q_j < q^*_j$ is larger than the number of physicians chosing q^*_j.[17] When comparing the average deviation $\overline{\nu}_j$ over patient types, the above differences appear to be supported. $\overline{\nu}_j$ varies between -0.14 and 0.45 for patient type 2. It fluctuates between -0.73 and -0.27 for patient type 1 and between -1.82 and -1.23 for patient type 3.

Results 4 and 5 evidence that our experimental physicians in CAP behave like we conjectured. Patients are underserved in that subjects on average choose quantities of medical services q^{CAP} smaller than the patient's optimal quantity q^*. Again, physicians' choices are strongly influenced by patient types (Result 6). The number of optimal choices increases when the difference between q^*_j and $\hat{q}_j$ becomes smaller, i.e. when q^*_j approaches zero. Remember that physicians maximize their profit by chosing zero medical services.

[16] 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.

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

4.1.3 Comparison between FFS and CAP

Our third research question is concerned with comparing behavior of experimental physicians across treatments.

Average quantities provided in FFS are about 50 percent larger than in CAP (6.60 vs. 4.40, see Table 7). Almost the same holds for the median (7.00 vs. 5.00) whereas the standard deviation is only slightly larger in FFS (1.85 vs. 1.64). We first analyze the data from the patient's point of view. To this end, we compare $\overline{q}_j^{FFS}$ and $\overline{q}_j^{CAP}$.

Result 7. *Patients are provided with more medical services in FFS than in CAP.*

SUPPORT: Figure 3 and Table A.5 show that each patient on average is treated with more medical services in FFS than in CAP. This difference is highly significant ($p = 0.0000$, Mann-Whitney U test, two-sided). Comparing individual physicians' services per patient, the picture is only slightly different. Except for patients 1A and 3A[18], physicians provide patients with highly significantly larger quantities in FFS than in CAP (all $p \leq 0.0010$, Mann-Whitney U test, two-sided; see line II in Table A.4). Thus, a significantly higher number of patients are provided with significantly more medical services in FFS compared to CAP ($p = 0.007$, binomial test, two-sided).

The next result is concerned with decisions of the individual physician in the two treatments. We compare the quantity of medical services each physician provides averaged over the 15 different patients, i.e. $\overline{q}_i^{FFS}$ and $\overline{q}_i^{CAP}$.

Result 8. *Physicians in FFS provide larger quantities than physicians in CAP.*

SUPPORT: On average physicians in FFS provide services significantly larger than in CAP ($p = 0.000$, Mann-Whitney U test, two-sided). Furthermore, $\overline{\mu}_i \geq 0$ in FFS, and except for physician $i = 4$, $\overline{\mu}_i \leq 0$ in CAP (see Table A.1).

Results 7 and 8 support our conjecture on physicians' behavior across treatments. Patients receive much more medical services in FFS than in CAP. We now analyze the impact of q_j^* on physicians' behavior across treatments.

Result 9. *The patient's optimal quantity of medical services and its values differing with patient types influences physicians' decisions more decisively in CAP compared to FFS.*

SUPPORT: We first analyze physicians' choices with regard to q_j^* across treatments. See also Figure 2 for absolute frequencies of $q{^*}_{ij}^{FFS}$ and $q{^*}_{ij}^{CAP}$. We find that physicians in CAP choose the patient optimal quantity of medical services significantly more often than physicians in FFS do ($p = 0.014$, Mann-Whitney U test, two-sided).

When studying q_j^*–choices across treatments for each patient kh (decision j) separately, decisions are found to depend on patient types. In CAP, all patients of type 2 get a better treatment in that significantly more physicians chose q_j^* than in FFS (Fisher exact test, see line III in Table A.4). The same applies for patients of type 1 except for decision $j = 1$

[18] Here, $p = 0.2440$ for 1A and $p = 0.2339$ for 3A (Mann-Whitney U test, two-sided).

(patient 1A). In the latter case, physicians in FFS make significantly more q_j^*–choices. In fact, all 20 physicians provide $q_1^* = 5$, whereas in CAP only 15 of the 22 physicians behave accordingly. For patients of type 3 evidence is mixed; we find no significant difference for patients 3A, 3C, 3E (decisions $j = 11, 13, 15$). For patients 3B and 3D (decisions $j = 12, 14$), q_j^* is chosen significantly more often in FFS than in CAP.

The cross-treatment comparison provide evidence that physicians' choices with regard to q_j^* are highly influenced by the payment system as well as by patient types.

4.2 Profit and patient benefit

Our forth research question deals with physician's profit and patient benefit. We are interested to what extent and when experimental physicians take the patient benefit into account when making their quantity decisions. We have seen already that subjects do not behave in a completely self-interested manner. In this subsection, we will analyze this phenomenon in more detail.

Our previous results suggest that patient benefit $B(q_{ij})$ and physician's own profit $\pi(q_{ij})$ are major behavioral determinants in both treatments.[19] Recall that both variables are simultaneously determined by physicians' decisions. We also analyze patient benefit losses. We define a benefit loss $\psi(q_{ji})$ to occur for a patient whenever a physician deviates from choosing q_j^*, i.e. $\psi(q_{ji}) = |B(q_{ij}) - B(q_j^*)|$.

Table 8: Profit and patient benefit

		Mean	Median	SD	Total number of decisions
FFS	Profit $\pi(q_{ij})$	9.17	8.00	2.69	300
	Patient benefit $B(q_{ij})$	8.83	9.00	1.10	300
CAP	Profit $\pi(q_{ij})$	9.79	9.50	1.52	330
	Patient benefit $B(q_{ij})$	8.56	9.75	2.46	330

Physicians profit

Remember that in FFS the maximum profit $\pi(\hat{q}_j)$ is 8.00 (12.50, 8.20, 13.60, 13.00) for illness A (B, C, D, E); see Table 6. Choosing $\hat{q}_j$ for all j would have yielded them an average payoff of 11.1. In CAP, the maximum profit is 12.00 for all illnesses. Physicians in our experiment provided quantities of medical services such that they get an average profit of 9.17 in FFS and 9.79 in CAP (Table 8) .

Result 10. *Physicians's profits do not differ across treatments, although the variance is larger in FFS than in CAP.*

[19]This is also supported by the experimental physicians' statements in the post-experimental questionnaires. Six of the 42 physicians chose $q_{ij} = q_j^* \;\; \forall j = 1, \ldots, 15$. Explaining their behavior throughout the experiment they stated for example "I wanted to act as a good physician caring for their patients" or "The patient benefit should be maximal". 40 subjects reported the patient benefit to have influenced their decisions. 20 stated to weigh own profit relatively to the patient benefit.

SUPPORT: There is no evidence in the data that profits per physician averaged over patients differ in the two treatments ($p = 0.332$, Mann-Whitney U test, two-sided). The variance is significantly larger in FFS, however ($p = 0.000$, Mann-Whitney U test, two-sided). This finding is corroborated when averaging profits over illnesses. In FFS, mean profits vary from 7.22 in illness C to 11.29 in illness D. In CAP, profits fluctuate between 9.63 to 9.93 (Table 9).

Table 9: Pysicians' average profit $\overline{\pi}_h(q_j)$ per illness in FFS and CAP

Illness (h)	$\overline{\pi}_h^{CAP}$	$\overline{\pi}_h^{FFS}$
A	7.47	9.63
B	9.95	9.82
C	7.22	9.82
D	11.29	9.77
E	9.92	9.93

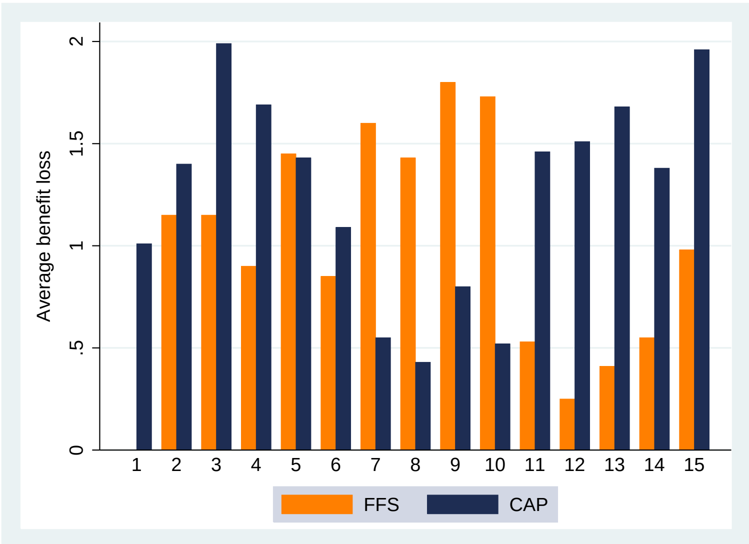
Patient benefit and patient benefit loss

In both treatments, the benefit optimum for patients of type 3 ($B_3(q_j^*)$) is 9.45. $B_1(q_j^*) = B_2(q_j^*) = 10$ (see Table 4). If physicians had always chosen the patient optimal quantity, patients would have received an average benefit of 9.82.

The actual data show average patient benefit $B(q_{ij})$ to be slightly larger in FFS (8.83) than in CAP (8.56). Further, average patient benefits determined by physician i vary between 7.52 and 9.82 in FFS and between 2.73 and 9.82 in CAP (see Table A.7). The data show no evidence that mean patient benefits $B(q_i)$ or the variances differ across treatments ($p = 0.504$, Mann-Whitney U test, two-sided).

Next we are concerned with differences in the benefit loss per patient across treatments.

Figure 5: Average benefit loss per patient



Result 11. *Benefit losses per patient differ across treatments; they depend on patient types.*

SUPPORT: Figure 5 contrasts the average benefit loss per patient across treatments. For

10 of the 15 patients ($kh = 1\text{A}, \ldots, 1\text{D}, 2\text{A}, 3\text{A}, \ldots, 3\text{E}$), the benefit loss is larger in CAP compared to FFS (see also Table A.6). For the remaining patients, the benefits loss is larger in FFS.

We find again that patient types matter. Test statistics of a two-sided Mann-Whitney U test yield that benefit losses differ significantly for each illness of patient type 2 (see row line III of Table A.4). In particular, losses are larger in FFS for patients 2B, 2C, 2D, 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. The losses of patient type 1 do not differ significantly except for patient A1 ($p = 0.009$) where no losses occur in FFS, and for A5 ($p = 0.062$). Also the losses of patient type 3 do not differ significantly, except for patients 3B ($p = 0.002$) und 3C ($p = 0.050$) ($j = 12$).

Result 11 suggests that for patients in need of a small quantity of medical services like patients of type 2, on average a smaller benefit loss results when physicians are paid by CAP. On the contrary, patients in need of a larger quantity of medical services, like patients of types 1 and 3, incur a smaller loss under a FFS system.

4.3 Tradeoffs between profit and patient benefit

Our last research question involves analysing the tradeoffs between own profit and patient benefit the experimental physicians are faced with. According to our parameters, 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 analyse whether behavior varies in the two payment systems and whether subjects differ in their choices with regard to tradeoffs. Figures A.1 and A.2 plot patient benefit against physician's profit, show allocations on the pareto frontier as well as the frequency of physician' decisions.

We classify the patient-benefit/profit combinations determined by physicians' quantity decisions for each patient in order to compare the relation between patient benefit and physician's profit. Choices of the profit-maximizing quantity $\hat{q}_j$ implying the combination $(B(\hat{q}_j), \pi(\hat{q}_j))$ constitute the first category PROMAX. The second category PATMAX consists of q^*_j-choices (maximizing the patient's benefit) involving the combination $(B(q^*_j), \pi(q^*_j))$. Table A.8 shows patient-benefit/profit combinations for both categories in treatments FFS and CAP. The third category PARETO comprises choices entailing patient-benefit/profit combinations located on the pareto frontier other than q^*_j- and $\hat{q}_j$-choices. The remaining choices constitute the last category OTHER.

We further characterize physician's decisions according to the slope between $(B(\hat{q}_j), \pi(\hat{q}_j))$ and $(B(q^*_j), \pi(q^*_j))$ illustrating the tradeoff size for each patient. We distinguish two cases, a *flat* and a *steep* slope. We define a slope as flat if the absolute value of the slope is smaller than 1 ($|\text{slope}| < 1$), i.e. the patient's marginal benefit is larger than the physician's marginal profit loss. A slope is defined as steep if its absolute value is larger than 1 ($|\text{slope}| > 1$), i.e. the physician's marginal profit is larger than the patient's marginal benefit loss. In other words, the flatter the slope the larger is the gain in additional benefit for the patient and the steeper the slope, the larger is the loss in the physician's profit necessary to increase a patient's benefit.

Tradeoffs in FFS

Table A.9 shows relative frequencies of physicians' choices for all 15 patient in both treatments in the above categories PROMAX, PATMAX, PARETO and OTHER. Only 4.60% of all choices did not entail a pareto-efficient patient-benefit/profit combination. Recall that $j = 1$ (patient 1A) does not imply a tradeoff between physician's profit and patient benefit as maximum profit and maximum patient benefit coincide at $q_j^* = 5$. All subjects chose this combination; see the parameters for patient 1A in Figure A.1.

Flat slope. For 6 of the 15 patients a flat slope exists (decisions $j = 3, 11, \ldots, 15$). Here, on average only 4.17% of the physicians belong to category PROMAX (see Table A.8). For 4 of these 6 patients ($j = 3, 12, 13, 14$) no physician chose $\hat{q}_j$. A mean percentage of 45.82% provide the maximum benefit to the patient (PATMAX). Thus, a considerable proportion of physicians was willing to increase the abstract patient's benefit provided giving up own profit was not sizeably large. In particular, this holds for patients 3B, 3C, and 3D ($j = 12, 13, 14$) where 61.66% of physicians chose q_j^*.

Steep slope. Of the 8 patients ($j = 2, 4, \ldots, 10$) with a steep slope, 24.38% of the subjects chose $\hat{q}_j$ (category PROMAX) and 21,88% belonged to category PATMAX. The largest proportion (51.88%) of experimental physicians belonged to category PARETO for 6 of 8 patients ($j = 2, 4, 5, 7, 8, 9$) .

Tradeoffs in CAP

In CAP, the relation between $(B(\hat{q}_j), \pi(\hat{q}_j))$ and $(B(q_j^*), \pi(q_j^*))$ is somewhat more uniform. Here, only flat slopes occur with absolute values of 0.25 for decisions $j = 1, \ldots, 5$, 0.09 for decisions $j = 6, \ldots, 10$ and 0.52 for decisions $j = 11, \ldots, 15$ (Table A.8). That is, when deciding for patients of type 3, physicians need to give up the largest amount of own profit to increase patient's benefit. Here, only 34.55% of the physicians belong to category PATMAX, yet 62.73% belong to category PARETO. Thus, for these patients the largest proportion of subjects belong to category PARETO. On the contrary, for patients of type 1 and 2 the largest proportion of physicians belong to category PATMAX; i.e. averaged over the five patients per type 66.36% and 83.64% chose q_j^* respectively.

The data in FFS and CAP suggest that the majority of physicians is willing to forego own profit only to a certain extent in order to increase patients' benefit. Furthermore, subjects are heterogeneous in their willingnwss to trade off own profit and patient benefit.

5 Conclusion

Our experimental results show that the payment system does influence experimental physicians' provision behavior. In particular, in FFS more medical services are provided compared to CAP. This is in line with theoretical findings (see e.g. Ellis & McGuire 1986, Schneider & Mathios 2006). We also found that patients with a low need for medical services suffer a lower benefit gain in CAP whereas patients with a high need for medical services gain a higher benefit in FFS. Our results show that the patient benefit has crucially influenced experimental physicians' decisions on the quantity of medical services. We conclude that we were able to indirectly include real patients into our experiment by donating the monetary patient benefit to a welfare organization caring for real patients. Our design thus elicited benevolent behavior towards the abstract patient in our experiment.

We are aware that a real physician-patient interaction cannot be modeled in each facet in a laboratory experiment. But from our results we conclude that our experimental setup is appropriate to investigate the influence of institutional factors on phyisicans' provision behavior. It may be premature to generalize the experimental results and to draw inferences about real world phyisican behavior. Nonetheless, we think that the present experimental investigation marks a first step to use the tool of laboratory experiments, as it provides an isolated and controlled analysis of factors influencing physicians' behavior. Making experiments more realistic by introducing uncertainty about the impact of medical treatments and patients' health status, demand side effects through patients, monitoring mechanisms, primary care physicians and specialists is an important challenge for future research.

References

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.

Camerer, C. (2003), *Behavioral Game Theory*, Princeton (NJ), Princeton University Press.

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

Choné, P. & Ma, C. (2006), Asymmetric Information from Physician Agency: Optimal Payment and Health Care Quantity. Working Paper, University of Boston.

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

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

Dumont, E., Fortin, B., Jacquement, N. & Shearer, B. (2007), Physicians' Multitasking and Incentives: Empirical Evidence from a Natural Experiment. Working Paper, Laval University.

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

Eisenberg, J. (1986), *Doctors' Decisions and the Cost of Medical Care*, Ann Arbor (MI), Health Administration Press.

Ellis, R. & McGuire, T. (1986), 'Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply', *Journal of Health Economics* **5**, 129–151.

Ellis, R. & McGuire, T. (1990), 'Optimal Payment Systems for Health Services', *Journal of Health Economics* **9**, 375–396.

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

Fan, C., Chen, K. & Kan, K. (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., Gächter, S. & Fehr, E. (2001), 'Are People Conditionally Cooperative? Evidence from a Public Goods Experiment', *Economics Letters* **71**, 397–404.

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

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.

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

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

Iversen, T. & Lurås, H. (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.

Jennison, K. & Ellis, R. (1987), Comparison of Psychatric Service Utilization in a Single Group Practice, *in* McGuire & Scheffler, eds, 'The Economics of Menatal Health Services: Advances in Health Economics and Health Services Research', Greenwich (CT), JAI Press.

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

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

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

Ma, C. (2004), Altruism and Incentives in Public and Private Health Care. Working Paper, University of Boston.

Ma, C. & McGuire, T. (1997), 'Optimal Health Insurance and Provider Payment', *American Economic Review* **87**, 685–704.

McGuire, T. (2000), Physician Agency, *in* Cuyler & Newhouse, eds, 'Handbook of Health Economics, Vol. 1 A', Amsterdam, North-Holland.

McGuire, T. & Pauly, M. (1991), 'Physician response to fee changes with multiple payers', *Journal of Health Economics* **10**, 385–410.

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

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

Schneider, H. & Mathios, A. (2006), 'Principal Agency Theory and Health Care Utilization', *Economic Inquiry* **44**, 429–441.

Scott, A. & Hall, J. (1995), 'Evaluating the Effects of GP Remuneration: Problems and Prospects', *Health Policy* **31**, 183–195.

Selten, R., Mitzkewitz, M. & Uhlich, G. (1997), 'Duopoly Strategies Programmed by Experienced Players', *Econometrica* **65**, 517–555.

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

A Appendix

A.1 Instructions (*translated from German*)

General Information

In the following economic experiment you will make a couple of decisions. If you follow the instructions carefully, you can (depending on your decisions) earn a considerable amount of money. Thus, it is important to read the instructions carefully.

You decide anonymously in your cubicles at your computer screens. During the experiment you are not allowed to talk with other participants. Whenever you have a question, please indicate it by raising your hand. Your question will be answered in private. If you disregard these rules you can be expelled from the experiment without receiving any payment.

Within the experiment all amounts of money are stated in Taler. At the end of the experiment your earnings from the experiment will be transferred at a rate of 1 Taler = 0,05 €.

Your decisions in the experiment

During the entire experiment you (like all other participants) are in the role of a physician deciding on how 15 patients should be treated, i.e. you decide on the **quantity** of medical services you want to provide per patient.

You decide at your computer screens. Here, subsequently five different illnesses – A, B, C, D and E – of three different types of patients – 1, 2 and 3 – occur. For each patient you can decide between $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ or 10 medical services to provide.

Instructions in treatment FFS

Your remuneration is as follows: For each **quantity** of medical services a different PAYMENT is assigned. The PAYMENT increases with the **quantity** of medical services.

Besides your PAYMENT you determine your COSTS while deciding on the **quantity** of medical services. COSTS increase with increasing **quantity** provided. Your PROFIT is calculated by subtracting your COSTS from your PAYMENT.

Further, from each quantity of medical services provided the patient gains a certain benefit, the PATIENT BENEFIT. That means, with your decision on the **quantity** of medical services you determine both your own PROFIT and the PATIENT BENEFIT. An illustrative example is given on the following screen.

Beschwerdebild B / Patient 1

Medizinische Leistungen	Anzahl	Ihre VERGÜTUNG (in Talern)	Ihre KOSTEN (in Talern)	Ihr VERDIENST (in Talern)	PATIENTENNUTZEN (in Talern)
keine	0	0.00	0.00	0.00	0.00
Leistung B1	1	1.00	0.10	0.90	0.75
Leistung B1 / Leistung B2	2	2.40	0.40	2.00	1.50
Leistung B1 / Leistung B2 / Leistung B3	3	3.50	0.90	2.60	2.00
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4	4	8.00	1.60	6.40	7.00
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5	5	8.40	2.50	5.90	10.00
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5 / Leistung B6	6	9.40	3.60	5.80	9.50
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5 / Leistung B6 / Leistung B7	7	16.00	4.90	11.10	9.00
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5 / Leistung B6 / Leistung B7 / Leistung B8	8	18.00	6.40	11.60	8.50
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5 / Leistung B6 / Leistung B7 / Leistung B8 / Leistung B9	9	20.00	8.10	11.90	8.00
Leistung B1 / Leistung B2 / Leistung B3 / Leistung B4 / Leistung B5 / Leistung B6 / Leistung B7 / Leistung B8 / Leistung B9 / Leistung B10	10	22.50	10.00	12.50	7.50

Bitte wählen Sie die Anzahl medizinischer Leistungen, die Sie zu obigem Beschwerdebild als notwendig erachten.

Ihre Entscheidung:

OK

Instructions in treatment CAP

Your remuneration is as follows: For each **quantity** of medical services you receive the same PAYMENT.

Besides your PAYMENT you determine your COSTS while deciding on the **quantity** of medical services. COSTS increase with increasing **quantity** provided. Your PROFIT is calculated by subtracting your COSTS from your PAYMENT. An illustrative example is given on the following screen.

Beschwerdebild C / Patient 1

Medizinische Leistungen	Anzahl	Ihre VERGÜTUNG (in Talern)	Ihre KOSTEN (in Talern)	Ihr VERDIENST (in Talern)	PATIENTENNUTZEN (in Talern)
keine	0	12.00	0.00	12.00	0.00
Leistung C1	1	12.00	0.10	11.90	0.75
Leistung C1 / Leistung C2	2	12.00	0.40	11.60	1.50
Leistung C1 / Leistung C2 / Leistung C3	3	12.00	0.90	11.10	2.00
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4	4	12.00	1.60	10.40	7.00
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5	5	12.00	2.50	9.50	10.00
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5 / Leistung C6	6	12.00	3.60	8.40	9.50
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5 / Leistung C6 / Leistung C7	7	12.00	4.90	7.10	9.00
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5 / Leistung C6 / Leistung C7 / Leistung C8	8	12.00	6.40	5.60	8.50
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5 / Leistung C6 / Leistung C7 / Leistung C8 / Leistung C9	9	12.00	8.10	3.90	8.00
Leistung C1 / Leistung C2 / Leistung C3 / Leistung C4 / Leistung C5 / Leistung C6 / Leistung C7 / Leistung C8 / Leistung C9 / Leistung C10	10	12.00	10.00	2.00	7.50

Bitte wählen Sie die Anzahl medizinischer Leistungen, die Sie zu obigem Beschwerdebild als notwendig erachten.

Ihre Entscheidung:

OK

You decide on the **quantity** of medical services at your computer screen by typing in a number between 0 and 10 into the field “Your Decision”.

There are no real patients participating in this experiment; patients are rather abstract. But the PATIENT BENEFIT an abstract patient recieves through your quantity decisions will be beneficial for a real patient. The summed up amount of all 15 PATIENT BENEFITs determined by your decsions will be transferred to the charitably organization *Christoffel Blindenmission Deutschland e.V., 64625 Bensheim* in order to support an ophthalmic hospital where patients with cataract are treated.

Earnings in the experiment

After your 15 decisions, your overall earnings will be calculated by summing up your PROFITs and transferring them from Taler into Euro.

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

The transmission will be done by the experimenter and a control person. The control person inscribes the amount of money resulting from summing up overall PATIENT BENEFITs of all subjects into a crossed check. This check is issued to the *Christoffel Blindenmission* and will be put into an envelope addressed to this charity. The envelope will be thrown into the nearest mail box.

After all subjects took their decisions, one participant is randomly assigned the role of the control person. The control person receives an additional payment of 4 €. By signing a document the control person states that the procedure described here was actually carried out.

In the following we would like to ask you to answer some questions familiarizing you with the decisions in the experiment.

After your decisison in the experiment you are asked to complete some questions at your screen.

A.2 Data and statistics

Table A.1: Mean quantity ($\overline{q}_i$) and mean difference ($\overline{\mu}_i$) per physician i

	FFS		CAP	
i	$\overline{q}_i$	$\overline{\mu}_i$	$\overline{q}_i$	$\overline{\mu}_i$
1	6.40	1.40	4.20	-0.80
2	7.73	2.73	4.27	-0.73
3	5.00	0.00	4.80	-0.20
4	5.00	0.00	5.13	0.13
5	7.27	2.27	2.13	-2.87
6	6.40	1.40	5.00	0.00
7	7.13	2.13	4.07	-0.93
8	8.27	3.27	4.33	-0.67
9	6.07	1.07	4.07	-0.93
10	7.67	2.67	5.00	0.00
11	7.47	2.47	4.93	-0.07
12	6.93	1.93	4.93	-0.07
13	6.13	1.13	2.40	-2.60
14	6.27	1.27	5.00	0.00
15	8.53	3.53	4.00	-1.00
16	6.67	1.67	4.47	-0.53
17	5.00	0.00	3.40	-1.60
18	5.73	0.73	4.53	-0.47
19	7.00	2.00	6.00	1.00
20	5.33	0.33	4.67	-0.33
21			5.00	0.00
22			4.47	-0.53

Table A.2: Test statistics of a Fisher-Pitman permutation test for quantity decisions per physician i

	Treatment	
i	FFS	CAP
1	0.00195313	0.0976563
2	0.00512695	0.0703125
3	1.00000000	0.2500000
4	1.00000000	1.0000000
5	0.00012207	0.0001221
6	0.01562500	1.0000000
7	0.00122070	0.0078125
8	0.00085449	0.0625000
9	0.01171875	0.0195313
10	0.00195313	1.0000000
11	0.00512695	1.0000000
12	0.00390625	1.0000000
13	0.02343750	0.0009766
14	0.00781250	1.0000000
15	0.00085449	0.0019531
16	0.02539063	0.5000000
17	1.00000000	0.0136719
18	0.02734375	0.0625000
19	0.00341797	0.1829834
20	0.36914063	0.0625000
21		1.0000000
22		0.0625000

Table A.3: Number of patient optimal choices (q^*) and non-optimal choices ($\neg q^*$)

			Decision j (Patient kh)														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
			(1A)	(1B)	(1C)	(1D)	(1E)	(2A)	(2B)	(2C)	(2D)	(2E)	(3A)	(3B)	(3C)	(3D)	(3E)
FFS	q^*		20	5	6	4	3	5	2	5	3	8	8	15	10	12	4
	$\neg q^*$	total	0	15	14	16	17	15	18	15	17	12	12	5	10	8	16
		underprovision	0	0	1	0	0	0	0	0	0	0	11	2	2	2	3
		overprovision	0	15	13	16	17	15	18	15	17	12	1	3	8	6	13
CAP	q^*		15	15	15	13	15	14	18	20	20	20	9	7	8	8	6
	$\neg q^*$	total	7	7	7	9	7	8	4	2	2	2	13	15	14	14	16
		underprovision	5	6	7	7	6	2	1	1	2	1	13	15	13	14	16
		overprovision	2	1	0	2	1	6	3	1	0	1	0	0	1	0	0

Table A.4: Test statistics of two-sided non-parametric tests per patient

	Test; Variable(s); Scope	Decision j (Patient kh)														
		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^*, $\neg q_j^*$; within 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^*, $\neg q_j^*$; within 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_j; 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; $B(q_j)$; 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

Table A.5: Descriptive statistics on quantity q_j and difference ν_j per patient

		Decision j (Patient kh)														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		(1A)	(1B)	(1C)	(1D)	(1E)	(2A)	(2B)	(2C)	(2D)	(2E)	(3A)	(3B)	(3C)	(3D)	(3E)
FFS	$\overline{q}_j$	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 (q_j)	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
	$\overline{\nu}_j$	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 (ν_j)	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
	N	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
CAP	$\overline{q}_j$	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 (q_j)	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
	$\overline{\nu}_j$	-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 (ν_j)	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
	N	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22

Table A.6: Descriptive statistics on patient benefit $B(q_{ij})$ and benefit loss $\psi(q_{ij})$ per patient

			Decision j (Patient kh)														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
			(1A)	(1B)	(1C)	(1D)	(1E)	(2A)	(2B)	(2C)	(2D)	(2E)	(3A)	(3B)	(3C)	(3D)	(3E)
FFS	$B(q_{ij})$	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	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	$B(q_{ij})$	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
		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
	$\psi(q_{ij})$	Mean	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

Table A.7: 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.8: Patient-benefit/profit combinations in categories PATMAX and PROMAX for treatments FFS and CAP

	Category		Decision j (Patient kh)														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
			(1A)	(1B)	(1C)	(1D)	(1E)	(2A)	(2B)	(2C)	(2D)	(2E)	(3A)	(3B)	(3C)	(3D)	(3E)
FFS	PROMAX	$\pi(\hat{q}_j)$	8.00	12.50	8.30	13.60	13.00	8.00	12.50	8.30	13.60	13.00	8.00	12.50	8.30	13.60	13.00
		$B(\hat{q}_j)$	10.00	7.50	7.50	7.50	7.50	9.00	6.50	6.50	6.50	6.50	7.75	3.00	3.00	3.00	3.00
	PATMAX	$\pi(q_j^*)$	8.00	5.90	6.50	5.50	5.10	4.20	2.60	4.50	5.10	5.10	7.20	11.10	7.70	12.00	7.40
		$B(q_j^*)$	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	9.45	9.45	9.45	9.45	9.45
		Slope	0.00	-2.64	-0.72	-3.00	-3.16	-3.80	-2.83	-1.09	-2.43	-2.26	-0.47	-0.22	-0.09	-0.25	-0.87
CAP	PROMAX	$\pi(\hat{q}_j)$	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00
		$B(\hat{q}_j)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	PATMAX	$\pi(q_j^*)$	9.50	9.50	9.50	9.50	9.50	11.10	11.10	11.10	11.10	11.10	7.10	7.10	7.10	7.10	7.10
		$B(q_j^*)$	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	9.45	9.45	9.45	9.45	9.45
		Slope	-0.25	-0.25	-0.25	-0.25	-0.25	-0.09	-0.09	-0.09	-0.09	-0.09	-0.52	-0.52	-0.52	-0.52	-0.52

Table A.9: Relative frequencies of choices sorted by categories

		Decision j (Patient kh)														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		(1A)	(1B)	(1C)	(1D)	(1E)	(2A)	(2B)	(2C)	(2D)	(2E)	(3A)	(3B)	(3C)	(3D)	(3E)
FFS	PROMAX	1.00‡	0.25	0.00	0.15	0.30	0.65	0.15	0.10	0.10	0.25	0.20	0.00	0.00	0.00	0.05
	PATMAX	1.00‡	0.25	0.30	0.20	0.15	0.25	0.10	0.25	0.15	0.40	0.40	0.75	0.50	0.60	0.20
	PARETO	0.00	0.50	0.65	0.65	0.50	0.05	0.70	0.65	0.75	0.35	0.35	0.15	0.40	0.30	0.60
	OTHER	0.00	0.00	0.05	0.00	0.05	0.05	0.05	0.00	0.00	0.00	0.05	0.10	0.10	0.10	0.15
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
	PATMAX	0.68	0.68	0.68	0.59	0.68	0.64	0.82	0.91	0.91	0.91	0.41	0.32	0.36	0.36	0.27
	PARETO	0.23	0.27	0.32	0.27	0.23	0.09	0.05	0.05	0.09	0.00	0.55	0.68	0.59	0.64	0.68
	OTHER	0.09	0.05	0.00	0.09	0.05	0.27	0.14	0.05	0.00	0.05	0.00	0.00	0.05	0.00	0.00

‡ *Note that for the patient 1A (decision $j = 1$) in FFS* $(\pi(\hat{q}_j), B(\hat{q}_j)) = (\pi(q_j^*), B(q_j^*))$.

Figure A.1: Pareto frontiers FFS

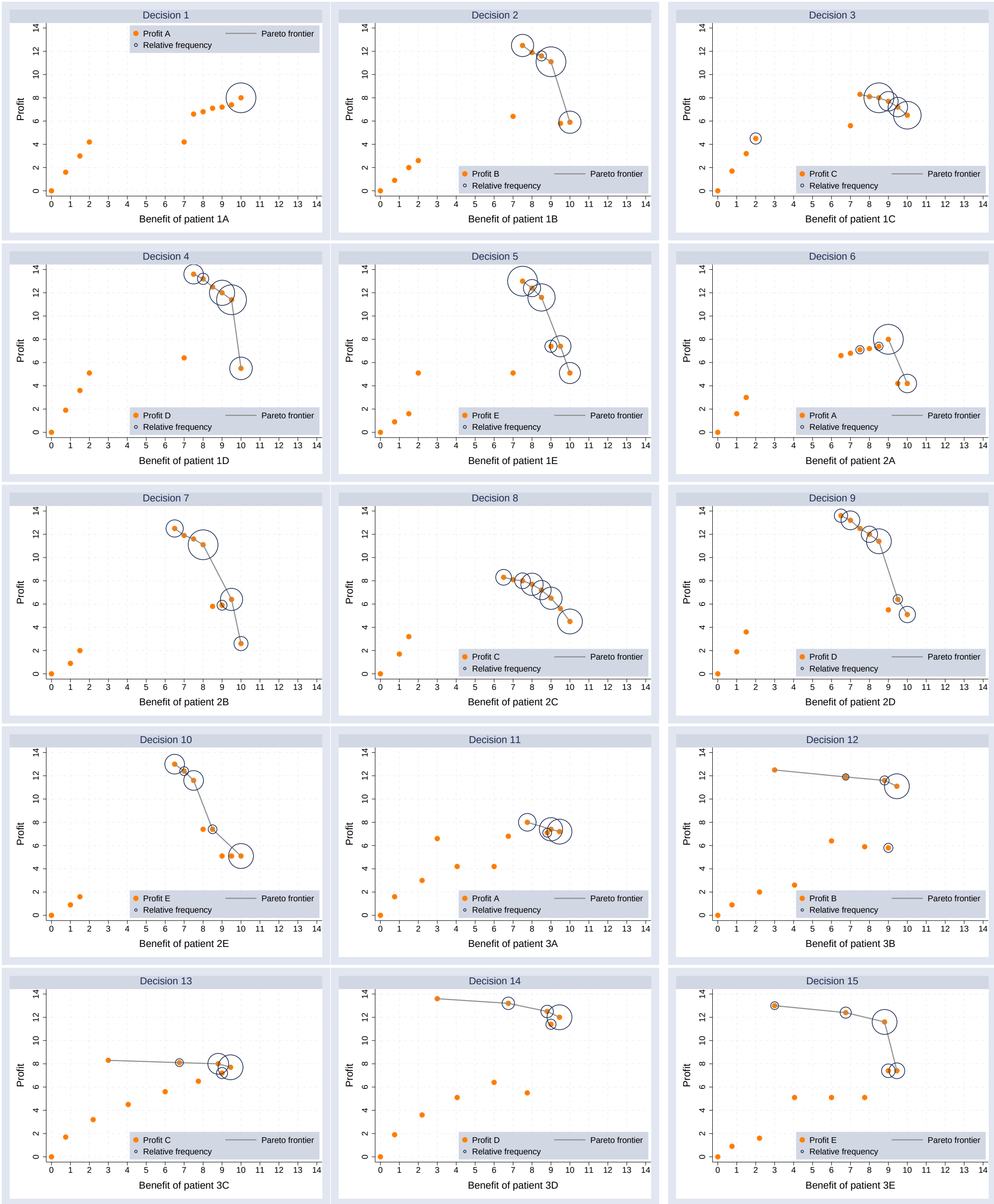


Figure A.2: Pareto frontiers CAP