The Expert-Client Information Problem

Fridolin E. Marty
University of Berne*

June 1999

Abstract

We consider a regulated market for credence goods where prices are fixed. Not knowing about the exact service she actually needs, the client has to rely on the advice of the expert. There are two types of experts: experts never cheating and opportunistic experts taking advantage of the information asymmetry. We compute Bayesian-Nash equilibria in pure and mixed strategies. The rejection strategy of the clients and the honest colleagues may prevent the cheating experts from always recommending a high price service. We discuss price setting strategies which reduce, with minimal welfare cost, the profits of the fraudulent experts as well as the amount of fraud. This kind of price setting is in line with a modified notion of an incentive compatible compensation scheme which is at the heart of current debates.

Keywords: Credence Goods, Expert, Fraud, Bayesian-Nash equilibrium.
Journal of Economic Literature Classification Numbers: C72, D82, I10

*Department of Economics, University of Berne, Gesellschaftsstrasse 49, CH-3012 Berne, Switzerland; email: fridolin.marty@vwi.unibe.ch, homepage: http://www-vwi.unibe.ch/theory/leb_fm.html.
I would like to thank Winand Emons, Mark de Snaijer, Gabrielle Wanzenried, and Martin Werner for helpful discussions and suggestions. Financial support from the Schweizerischer Nationalfonds is gratefully acknowledged.

1 Introduction

This article deals with services provided by experts. Since the specialization of labor increases, experts become more and more important. We speak of experts as suppliers when the consumers do not know the exact good or service they need. The expert advises his consumers with respect to their demand.[1] The literature refers to this kind of goods as credence goods. Even ex post the consumers are not able to determine whether they were served in their own interest or not. In consulting services, for example, problems may arise although the advice of the expert was accurate. The same holds true for health problems arising after a recent medical checkup. Moreover, overtreatment can hardly be detected. Once the problem is solved, the ex ante information is often no longer available. A filled tooth, for example, does not look rotten anymore regardless whether it was broken or not in the first place. Not surprisingly, the credence goods' literature questions whether non-fraudulent (market) equilibria exist within this information structure (e.g. Emons, 1997). The pioneering paper by Darby & Karni (1973) presupposes the existence of fraud in such markets when analyzing the *optimal amount of fraud*.

Due to this information asymmetry, credence goods markets are often regulated. In order to reduce the consumers' lack of information, experts have to use fee schedules which normally consist of the provided services and its respective prices. For several reasons, fee schedules which are incentive compatible are hardly feasible and sometimes even not desirable. First, the lack of information may be so severe that adequate prices cannot prevent overcharging. This is the case for experts' services that cannot be observed. Second, overtreatment is still possible as long as an observable treatment consists of many services which can be billed separately. Finally, incentive compatible prices are often very expensive for the consumers, since these prices would involve very high experts' incomes.

Therefore, we analyze the expert-client relationship within a fixed price setup identifying feasible prices which are also desirable in terms of equity. Furthermore, we study different information structures and introduce two types of expert allowing us to study their payoffs and to introduce a modified notion of an incentive compatible fee

[1] In order to distinguish expert (seller) and client (consumer) I use the male pronoun for experts and the female for customers.

schedule. This fee schedule, balancing the earnings of the two types of expert, reduces the incentive for fraudulent behavior.

Specifically, we consider the situation between a seller and a consumer within a regulated market where prices are fixed. The client wants an expert to solve her problem. The client does not know about the severity of her problem and the type of expert she is visiting. The expert states a diagnosis and proposes a problem solving strategy which can either be a treatment or an advice. The consumer decides to accept or reject the expert's proposal. When the proposal is accepted, the expert performs the service and the consumer pays for it. In case of rejection, the client consults another expert. Since we set up our model as a two-period-game and the consumer wants the problem to be solved, the advice of the second period expert is always accepted.

Two kinds of expert offer services to the consumers: The first type of expert fully acts in the interest of his clients. The second type, however, maximizes profits regardless of the consumers' demands. That is, he cheats whenever this is profitable for him. Therefore, honest experts are 'pathologically' honest, since they refuse to cheat even when there are incentives to do so. By contrast, dishonest experts are 'potentially' cheating since there are cases where they reveal honest behavior. This distinction follows Jaffee and Russell (1976) who investigate credit markets assuming honest and dishonest borrowers. In the controversy about physicians' salaries *the minority of physicians who just want to make money* are often blamed for spoiling the reputation of the whole profession.[2] Furthermore, anecdotal evidence suggests that consumers are sometimes cheated, while most of the experts do not take advantage of the information asymmetry.[3]

A crucial distinction of the information structure concerns the observability of the experts' service. Not surprisingly, the basic information problem increases, when the client is not able to observe the expert's service. In this case, the expert has the option to overcharge. Repair services performed when the client is absent or intellectual work like consulting services are examples for nonobservability. Nonobservable services which were not actually performed may be charged. In case of observability, however, overtreatment is the only possibility to cheat. The expert performs unnecessary treat-

[2] A physicians' representative quoted according to Tages-Anzeiger 15/5/98 (in italics).

[3] According to a representative of the Swiss health insurers, 10-15% of the physicians are 'black sheep' (SonntagsZeitung 13/12/98).

ment that is a source of inefficiency. Dentists, for instance, mainly perform observable services.

We analyze which types of equilibria result depending on the parameter environment and the information structure. All equilibria are unique. Theoretically, it is possible to set prices such that an efficient and non-fraudulent equilibrium results. The information problem would be solved in this equilibrium. The prices which sustain this equilibrium are, however, unfavorable for the consumers. Consumers are better off in an equilibrium containing fraud. That is, they suffer less from the information asymmetry in a fraudulent equilibrium.

In case of nonobservability, two types of fraudulent equilibrium arise. The first one is efficient. Dishonest experts overcharge and the consumers accept every recommendation. Accordingly, the switching costs are saved and no unnecessary treatment is performed. This fraudulent but efficient equilibrium in which customers never reject, however, generates a large income gap between the two types of expert. In order to reduce the fraudulent expert's payoff, consumers must occasionally reject an advice. Seeing a second expert, however, involves switching costs which could be avoided if the consumers always accepted. Fortunately, the efficiency loss is borne by the cheating experts, since their income significantly decreases.

In case of observability, fraudulent equilibria are never efficient due to the unnecessary treatment. Moreover, the equilibrium with rejection may be more efficient than the one without. On the other hand, the non-fraudulent and efficient equilibrium is more appealing in terms of equity since the consumers are not always worse off compared to the fraudulent equilibrium.

By studying the profits of the two types of expert, we identify prices which redistribute the profits among the two types. These prices establish a fee schedule which does not allow to take too much advantage from fraudulent behavior. Income balancing across the two types of expert lowers the incentive for fraudulent behavior. Although we do not endogenize the choice of type, income balancing is nevertheless important in order to prevent the honest type of expert to become fraudulent. Furthermore, such a fee schedule does not attract fraudulent types of expert from other areas considering a model with different fee schedules in force and mobile experts. Last but not least, honest experts who do their job satisfactorily should, according to common sense, not earn less than their cheating colleagues.

We discuss two price setting strategies for the regulator. Both of them assure that the equilibrium in mixed strategies is played where the clients occasionally reject an advice. First, the regulator may minimize the efficiency loss. Second, the regulator may minimize the ratio of the efficiency loss and the amount of redistribution. The latter assures that the redistribution is as cheap as possible.

There are two papers which use a similar setup. The first by Pitchik and Schotter (1987) describes a mixed strategy equilibrium in an expert-customer one shot game when prices are fixed. The expert randomizes between reporting either truthfully or not, whereas the customer randomizes between acceptance and rejection of a treatment recommendation. The authors specify the payoffs for each strategy ex ante such that a Nash equilibrium in mixed strategies exists. In contrast, we specify the conditions for mixed strategies to constitute an equilibrium. Whereas Pitchik and Schotter (1987) focus on the frequency of fraud, we are interested in welfare and in the distribution of payoffs. Pitchik and Schotter (1987) do not explicitly model the information structure. The incentive for dishonest behavior emerges from the larger profit for selling a major service compared to a minor one given that the latter is needed. In our paper, we distinguish between observable and nonobservable treatment, which directly influences the profit of dishonest behavior. This allows us to analyze the effect of observability on the resulting equilibria.

Contrary to our setup, Pitchik and Schotter (1987) only consider one type of expert who maximizes profits. In order to prevent the experts from always cheating, they specify an exogenous outside option for the customers. In our paper, the outside option is given by the fraction of honest experts. Accordingly, our equilibria degenerate into a pure strategy equilibrium when the fraction of the honest type converges to zero. Nevertheless, the results of Pitchik and Schotter (1987) could easily be mimicked within our setup with minor modifications.

The second paper by Wolinsky (1995) models equilibrium in a two period expert-customer game with an endogenous price setting. The consumers offer prices which can be turned down by the experts. Wolinsky (1995), assuming nonobservability, identifies several equilibria depending on the search costs. In case of high search costs, all customers are served for a price equal to marginal cost of the expensive problem in the first period. The expert always rejects a lower price in this unique equilibrium. Two kinds of equilibria exist for low search costs: One equilibrium is the same as before. The other equilibrium is interior in the sense that the expert mixes between accepting and rejecting a lower price offer. The customer always sees another expert if the expert rejects. This equilibrium is not unique since there are two probabilities of rejecting which satisfy the equilibrium conditions. Wolinsky (1995) identifies a mark-up over cost embodied in the prices of the small services despite intense competition.

In our paper, competition is only present in the form that experts facing a rejection lose clients. Price competition is, therefore, excluded by assumption. Accordingly, our model fits best in markets where competition among producers is poor. We have in mind the Swiss market for physicians' health services, for instance, where all prices are set by the government and physicians are not allowed to advertise.

Although prices are fixed in our model and two types of expert are introduced, our equilibria are similar to Wolinsky's. The main difference is that our consumers do not have to reject each major service recommendation in order to discipline the experts. It suffices to reject with a positive probability in order to reduce the amount of fraud. The honest experts prevent their fraudulent colleagues from playing a strategy of constant cheating.

Whereas Wolinsky's model fits in a competitive environment, our setup focuses on regulated markets with fixed prices set by a regulator or as a result of a bargaining process like in the Swiss market for physicians' services or dentists' services. Moreover, due to the distinction of two types of expert, we are able to raise the weakened notion of an incentive compatible compensation scheme. This fee schedule redistributes across the two types and cannot be exploited by opportunistic experts. The latter aim plays a crucial role in the bargaining process for compensation schemes.

Disciplining experts when the fee schedule is not perfectly incentive compatible is possible because the consumers look for a second opinion. In practice, health insurers promote seeking second opinions for certain operations. The usefulness of this policy is confirmed by our findings. Second opinions may save costs for insurance companies and customers although they may be costly in terms of efficiency.

The paper is organized as follows: In section 2, we present the model and its solution. Section 3 introduces the classical trade-off between efficiency and equity.

Two price setting strategies are suggested in order to optimize this trade-off. Section 4 analyzes the model for observable service comparing the results of the former sections. The last section concludes. Proofs and the normal form of the game are relegated to the appendix.

2 The model with nonobservable service

2.1 The basic model

We consider a market for credence goods. A continuum of consumers with measure 1 has either a major or a minor problem. An exogenous fraction $w \in (0,1)$ of customers suffers from a major problem (H) whereas the fraction $(1-w)$ has a minor one (L). A customer knows that she has a problem but does not know how serious it is. When her problem is solved she obtains utility of B. A treatment of the major problem also solves the minor problem but not vice versa.

The market consists of experts with measure 1, who diagnose and repair the clients' problems. A fraction $g \in (0,1)$ of experts (type g) fully acts in the interest of their clients, so they will never cheat on their customers. In contrast, a fraction $(1-g)$ of experts (type b) behave opportunistically, i.e., they may sell a major service to customers who only suffer from a minor problem. We capture this strategic decision by $x \in [0,1]$, which stands for the probability of cheating. Accordingly, the goal for a type b expert is to maximize his profits regardless of the customer's needs. The type b experts may be interpreted in different ways: first, they can be considered as selfish persons who just want to make money on their job. Second, they may be regarded as persons subject to a special economic environment. Idle capacities (Marty 1998) or economic pressure due to a high mortgage are examples for situations in which the economic incentives may be stronger than the ethic demands.

The experts' marginal costs of the minor and the major problem equal c_L and c_H, respectively. We normalize the marginal cost of the minor problem, i.e., $c_L = 0$. Additionally, fixed costs are present so we have to specify the prices above the marginal costs such that on average the experts break even. Price and marginal cost of the diagnosis equal zero in order to exclude the distinction between diagnosis and treatment from strategic considerations. This distinction is sometimes difficult to make.

Furthermore, some experts like psychiatrists or consultants treat the problem just by diagnosing it. The prices for both problems are fixed at p_H and p_L respectively. The prices include the treatment and satisfy $p_L \leq p_H$.

Following Wolinsky (1995), we assume that the existence of a problem is both observable and verifiable. The type of service (H or L) is, however, not observable to customers.[4] That is, consumers neither know which service is needed nor do they observe which treatment is actually performed. This means that payments can be conditioned to the solution of a problem but not to the type of treatment. In addition, it implies that an expert might be induced to misrepresent a minor service as a major one. That is, nonobservable services allow for overcharging.

We consider a two period game. In the first period, a customer visits an expert who is either of type g or of type b. The customers have the option to reject a service. They never reject an L-service because this is the cheapest way to obtain utility B. Since the customers do not know the type of expert, they randomly reject a high service offer in period 1 with probability $y \in [0, 1]$. In case of rejection, they visit another expert which leads to switching costs of k for the consumers. Although they still do not know their own type of problem, they always follow the advice of an expert in the second period, since the game ends in the second period and we assume $B > k + p_H$.[5] The latter assumption ensures that the consumer maximizes her utility by accepting in the second period.

Contrary to Wolinsky (1995), we assume that the expert recognizes a second period customer. In health services, for example, it is easy for the physician to find out from the anamnesis of his client whether she asks for a first or a second opinion. Additionally, it is sometimes in the interest of the client to reveal her 'search age'. This is the case for an expensive diagnosis which can be used by the second expert. Furthermore, some diagnosis procedures are injurious to health. When X-ray pictures have been taken, e.g., the patient hardly hides it. Therefore, an expert of type b always offers the high price service H to a second opinion customer since he is sure that his advice will be accepted.

[4] Section 4 analyzes the case of observability.

[5] This assumption is quite plausible for the health market, where utility B is achieved when the consumer is healthy. The demand for health services is likely to be inelastic.

The time structure of the expert-client relationship is summarized as follows:

$t = 0:$	the consumer identifies a problem not knowing if it is a minor or a major one, and decides to see an expert
$t = 1.1:$	the expert proposes either a minor or a major treatment.
$t = 1.2:$	the consumer decides whether to accept the advice or to see another expert.
$t = 2.1:$	in case of rejection, the second opinion expert proposes a treatment.
$t = 2.2:$	the consumer accepts.

2.2 Strategies

We consider Bayesian-Nash equilibria of this game. The experts have the strategies $\widetilde{H}/H$, $\widetilde{L}/H$, $\widetilde{L}/L$, and $\widetilde{H}/L$ in both periods. The strategy $\widetilde{L}/H$ is to be read as follows: recommend service L to a customer who actually needs the service H. $\widetilde{L}/H$ is never chosen, because the problem is both observable and verifiable. Accordingly, all experts recommend a major service to a customer who has a major problem (strategy $\widetilde{H}/H$). In addition, an expert of type g always chooses $\widetilde{L}/L$ by assumption, i.e., he recommends a minor service to a customer with an L problem in both periods. A type b expert, on the contrary, chooses in the first period with probability x the strategy $\widetilde{H}/L$, i.e., he tries to sell with probability x a major service to a customer with an L problem, when the price for the H-service is higher than for the L-service. In the second period, type b experts sell a major service to all customers, because they accept it for sure. As mentioned earlier, customers always accept a minor service recommendation since it is the cheapest way to solve their problem. In the first period, a major service is rejected with probability y. The incentive to reject an H-service in the first period is the prospect of getting an L-service recommendation in the second period. This is only possible with a type g expert. That is, some consumers are unlucky enough to end up with a type b expert paying for the expensive service.

2.3 Bayesian-Nash equilibrium in mixed strategies

Let us first consider the equilibrium in mixed strategy of this game. In the first period, the customers reject a major service with probability y in order to set a type b expert indifferent between cheating and telling the truth when facing a client with a minor problem.

$$(1-y)\, p_H = p_L \tag{1}$$

The profit of type b experts who tell the truth is p_L. By cheating he needs to be lucky in order to earn p_H. His luck is dependent on the rejecting strategy (y) of his client. Accordingly, the customer randomizes over rejecting (y) and accepting $(1-y)$ with

$$y^* = \frac{p_H - p_L}{p_H} \tag{2}$$

The type b expert chooses to propose the wrong service with probability x such that the customer is indifferent between accepting and rejecting.

$$B - p_H = B - k - \left(g \cdot prob(L|\widetilde{H})\right)\ p_L - \left[1 - (g \cdot prob(L|\widetilde{H}))\right]\ p_H \tag{3}$$

where

$$prob(L|\widetilde{H}) = \frac{(1-w)(1-g)\, x}{(1-w)(1-g)\, x + w} \tag{4}$$

is the probability of suffering from a minor problem despite receiving an H-advice in the first period.

A customer who receives the advice of an H-service revises the prior probability of being with an honest expert from g to $wg \,/\, [(1-w)(1-g)\, x + w]$ $(< g)$. The only possibility to benefit from rejecting in the first period is to receive an L-advice in the second period. This happens if the customer meets an honest expert in the second period *and* she has a minor problem given that she received an H-advice in the first period. The probability of this joint event is $g \cdot prob(L|\widetilde{H})$. We obtain $prob(L|\widetilde{H})$ by considering the event of receiving an H-advice despite having a minor problem. Naturally, an H-advice is also obtained if the customer actually has a major problem. Therefore, the overall probability of receiving an H-advice is $(1-w)(1-g)\, x+w$, whereas the probability of receiving an H-advice with an L-problem is only $(1-w)(1-g)\, x$. Accordingly, having a minor problem and receiving an H-advice occurs with $prob(L|\widetilde{H})$.

When the customer always accepts the recommendation of the expert, she then takes the risk to be cheated in the first period. She avoids, however, the switching costs k and the risk of obtaining an H-advice again in the second period despite having rejected it in the first period. The latter is happening either because she ends up with a type b expert or because she actually has a major problem. A customer who always rejects the advice in the first period has to incur switching costs but is cheated less frequently, since there is a chance of meeting a type g expert in the second period.

It follows from (3) that the type b expert randomizes between cheating (x) and telling the truth $(1-x)$ according to

$$x^* = \frac{w\,k}{(1-w)\,(1-g)\,(g\,(p_H - p_L) - k)} \tag{5}$$

The two probabilities $y^* \in [0,1]$ and $x^* \in [0,1]$ constitute a Bayesian-Nash equilibrium when at least one probability is strictly higher than 0 or strictly lower than 1.

Proposition 1: *There exists a Bayesian-Nash equilibrium in strictly mixed strategies for*

$$0 < \gamma k := \frac{(1-g+wg)\,k}{g\,(1-g)(1-w)} < p_H - p_L, \tag{6}$$

where $\gamma\,k$ is the cut-off price difference. The type b experts choose strategy $\widetilde{H}/L$ with probability x^* and the clients reject an H-advice with probability y^*.

We observe some amount of fraud in the mixed strategy Nash equilibrium. Due to the switching costs $k > 0$, the equilibrium is inefficient. The efficiency loss increases with the frequency of rejections which is dependent on both the probability of rejection and the probability of cheating in the first period. Given the mixed equilibrium, the efficiency loss is especially large for a low mark-up of the minor service. The effect of a large price differential of the major and the minor treatment is equivocal since in this case the probability of rejection is large, but the probability of cheating is small.

2.4 Bayesian-Nash equilibria in pure strategies

2.4.1 The switching costs are strictly positive

For strictly positive switching costs k, we observe two different settings depending on the value of $\gamma\, k$.[6] When the parameter values are such that $0 < p_H - p_L < \gamma\, k$ we obtain a Nash equilibrium in pure strategies where the customers always accept and the type b experts always cheat. The price differential between the major and the minor treatment is too low to make it worthwhile rejecting an H-advice in the first period. Or, to put it differently, the switching costs are too high relative to the price differential. In this equilibrium, fraud is at its maximum value. This equilibrium, though very unpleasant for the customers, is efficient. The switching costs are saved and no unnecessary treatment is performed. Solely for distributional reasons this equilibrium is unfavorable for the clients.

A non-fraudulent Nash equilibrium in pure strategies exists for $p_H = p_L$. Here, the type b experts have a weakly dominant strategy not to cheat, and the customers always accept for any value of $\gamma\, k$. Apart from its non-fraudulent property, this equilibrium has the same properties like the one before: it is efficient but distributes the payoffs one-sided. The experts obtain their maximum profit, i.e., they receive the highest price for all services.

2.4.2 The switching costs are zero

When the switching costs are zero, there are two types of Nash equilibria in pure strategies exist. They are both efficient since the switching costs are the only cause of inefficiency. Accordingly, the parameter environment exclusively determines the payoffs and the strategies played. Two different types of equilibria emerge depending on the mark-up of the minor service.[7] The mark-up of the major service, $p_H - c_H$, is irrelevant for the equilibrium selection since the incentive to overcharge only depends on the relation between p_L and p_H.

When $p_L > 0$, there is aNash equilibrium in pure strategies. The customers always reject a major service advice in the first period. In the second period, they accept both

[6] The cut-off value $\gamma\, k$ is always strictly positive for strictly positive switching costs, and it is always zero for zero switching costs.

[7] Since c_L is normalized to zero, the mark-up of the minor service is equal to p_L.

recommendations. This rejection policy forces the experts to non-fraudulent behavior in the first period. The type b expert prefers to earn p_L to earn nothing. In the second period, the experts would always cheat, however, only customers with a major problem end up in the second period. Accordingly, no customer with the minor problem is cheated in this equilibrium.

For $p_L = 0$, we obtain another type of Nash equilibrium in pure strategies. Here, the expert does not earn anything by selling a minor treatment, i.e., the fixed costs cannot be covered by a minor service treatment. Therefore, the type b experts always try to sell a major service. The customers reject every major treatment advice in the first period. Nevertheless, they cannot avoid being cheated in the second period since there is a positive probability of being treated by a type b expert after all. The probability for this event is equal to $[(1-w)(1-g)^2]$.

In the following Proposition we summarize all types of equilibria in pure strategies.

Proposition 2: *We obtain four types of equilibria in pure strategies. All equilibria are unique.*

(i) For $0 < p_H - p_L < \gamma k$*, with* $x^* = 1$ *and* $y^* = 0$.

(ii) For $p_L = p_H$*, with* $x^* = 0$ *and* $y^* = 0$.[8]

(iii) For $0 = p_L < p_H$ *and* $k = 0$*, with* $x^* = 1$ *and* $y^* = 1$.

(iv) For $0 < p_L < p_H$ *and* $k = 0$*, with* $x^* = 0$ *and* $y^* = 1$.

In the following, we neglect the case $0 < \gamma k = p_H - p_L$. Here, the consumers' payoffs are equal independent of whether the equilibrium in pure or in mixed strategies is played. We obtain a Nash-equilibrium in pure strategies, when type b experts choose $x^* = 1$ and consumers choose $y^* = 0$, otherwise we obtain a Nash-equilibrium in strictly mixed strategies.

The efficiency properties of the equilibria in pure strategies and in mixed strategies, respectively, are summarized in the following Corollary:

Corollary 1: *All equilibria in pure strategies are efficient, because the switching costs are avoided or equal to zero. The equilibrium in mixed strategies, however, is inefficient.*

As our definition of k suggests, it is not plausible to assume zero switching costs. Even if an expert provides a diagnosis for free and the new expert does not need to

[8] Here, we neglect the equilibrium $x^* = 1$ and $y^* = 0$, which is actually equivalent to *(ii)*.

perform another diagnosis, searching another expert is costly.[9] Accordingly, we deal with equilibria with strictly positive switching costs in the subsequent sections.

Summarizing all the results at hand, Table 1 concludes this section presenting an overview of the different types of equilibria:

type of equilibrium	strategies	parameter environment	behavior of type b experts	efficiency
a	mixed	$0 < \gamma\, k < p_H - p_L$	fraudulent	inefficient
b	pure	$0 < p_H - p_L < \gamma\, k$	fraudulent	efficient
c	pure	$p_L = p_H$	non-fraudulent	efficient
d	pure	$0 = \gamma\, k = p_L < p_H$	fraudulent	efficient
e	pure	$0 = \gamma\, k < p_L < p_H$	non-fraudulent	efficient

Table 1: Properties of the different equilibria

2.5 Utility of the consumers

Customers' expected utility U_a in the equilibrium in mixed strategies is equal to:

$$U_a = B - p_L - \frac{g\, w\, (p_H - p_L)^2}{g\, (p_H - p_L) - k} \tag{7}$$

The expected utility in mixed strategies U_a, which always involves a certain amount of fraud, is lower than the expected utility in an equilibrium in pure strategies without fraud: $U_c = U_e = B - p_L - w\,(p_H - p_L)$ that we regard as an upper benchmark. According to Proposition (2.4.2), this benchmark utility is only reached in equilibrium for $p_H = p_L$ or for $k = 0$ and $p_L > 0$. The expected utility for the customer in the Nash equilibrium with pure strategies and fraud is

$$U_b = B - p_L - (1 - g + wg)\,(p_H - p_L) \tag{8}$$

which equals the expected utility in the mixed strategy equilibrium for the cut-off value $\overline{k} = (p_H - p_L)\,\gamma^{-1}$. When the switching costs are equal to $\overline{k}$, the customer is just indifferent between accepting and rejecting, given that the type b expert is always cheating.

[9] Pesendorfer and Wolinsky (1998) call this kind of service *not appropriable*.

- insert figure 1 and 2 about here -

Figure 1 displays the expected utility of a consumer as a function of the switching costs k. The utility is decreasing in k up to the cut-off value $\overline{k}$, because the mixed strategy is played with a certain amount of rejection. For switching costs larger than $\overline{k}$, the consumer plays the Nash equilibrium with pure strategies where she always accepts in the first period. Accordingly, her utility is independent of k.

Figure 2 shows the expected utility of a consumer as a function of p_L. The utility is falling in price. Again, there is a limit price $\overline{p}_L$ equal to $p_H - \gamma\, k$. If p_L is higher than $\overline{p}_L$, the consumer is playing the pure strategy equilibrium, and for lower values of p_L the mixed strategy equilibrium is played. At $p_L = p_H$, the utility is at its minimum, although the experts are honest.

2.6 Experts' profit

In the pure equilibrium without fraud, both types of expert earn $w(p_H - c_H) + (1 - w)\, p_L$ per client on the average. The earnings differ, however, when we consider the pure equilibrium with fraud and customers' acceptance. Whereas the type g expert remains at the same amount, the type b expert is able to increase his earnings up to $p_H - wc_H$ per client. He advises as if he only faced customers with an H-problem, but makes even more money because the marginal costs differ across services.

The equilibrium in mixed strategies equalizes to some extent the earnings per customer across all experts. Both types earn exactly the same expected amount per first period customer which is $[p_L(1 - wc_H/p_H)]$. Nevertheless, type b experts treat relatively fewer first period customers than type g experts, i.e., due to their recommendation policy the type b experts lose a fraction of $(1 - w)\, x\, y$ consumers. Therefore, type b experts have a higher turnover per client than type g experts. Not only do they treat fewer first period customers but they also bill a major service to customers who actually need a minor one. Furthermore, type b experts bill a high service to all second period customers, hence the turnover per second period customer is also higher for type b experts.[10] As a result, type b experts face a higher average income per client. Not

[10] Marty (1998) empirically shows that idle capacities and a high turnover per patient go hand in hand. Therefore, the type b experts can be seen as experts with idle capacities.

only do they overcharge some of the accepting first period customers, but also they can profit from second period customers with an L-problem who were with another type b expert in the first period. Nevertheless, playing the mixed equilibrium reduces the earnings of the fraudulent experts substantially. When a client rejects an H-advice, the cheating expert even lacks the small mark-up p_L which he could have earned for sure. This mark-up p_L is earned by the type g experts when the clients are lucky to end up with an honest expert in the second period. Contrary to their colleagues, type g experts are, therefore, better off by playing the mixed equilibrium as their number of clients is higher.[11]

- insert figure 3 about here -

Figure 3 displays the population payoffs of the type g and type b experts as a function of p_L. By population payoff we mean the expected sum of all possible payoffs generated by the respective type of expert in both periods. Up to the cut-off price $\overline{p}_L$, the mixed strategy equilibrium is played. For prices higher than p_L, experts and customers choose a pure strategy. The type b experts drastically profit from playing the pure strategy. For prices lower than $\overline{p}_L$, their payoff is increasing in p_L, and it is a little higher than the payoff of the type g experts due to their second period customers with a minor problem which are overcharged. At $\overline{p}_L$, however, their payoff jumps to a value significantly higher than the payoff of their type g colleagues. For higher prices than $\overline{p}_L$, the payoff of type b experts is independent of p_L since they only bill major services. The population payoff of the type g experts is an increasing function of p_L for all relevant prices. Contrary to their type b colleagues, the type g experts profit from a change from the equilibrium in pure strategies to the mixed strategy equilibrium at price $\overline{p}_L$. They treat a larger clientele in the mixed strategy equilibrium. The reason is that some of the customers who were cheated in the first period are lucky to meet a type g expert in period two. Due to a positive mark-up of the minor problem, the honest experts enjoy extra profits compared to the pure strategy equilibrium.

[11] This is unambiguously true at price $\overline{p}_L$ only.

2.7 Comparative statics

In the previous sections we saw that the distributional properties of the equilibrium in mixed strategies are very appealing. In this section we analyze the likelihood of such an equilibrium depending on the parameters. Additionally, we study how the amount of fraud behaves within our setup. For these reasons, we perform comparative statics exercises.

Equation (6) gives the condition for the existence of an equilibrium in mixed strategies. Equilibria in mixed strategies are more likely the lower the switching costs (k) and the larger the price differential between the major and the minor treatment ($p_H - p_L$). The switching costs can be interpreted as the 'price' of rejection whereas the price differential is the uncertain benefit of it. Furthermore, these equilibria occur with a higher probability when a major treatment is less likely. In this case, an H-advice induces the customers to revise their belief about the type of expert more strongly.[12] This is the intuitive effect which arise if the expert's suggestion strongly deviates from the customer's expectation. Accordingly, rejecting an H-treatment in the first period becomes more attractive.

Finally, the relation between the likelihood of an equilibrium in mixed strategies and the fraction of type g experts is equivocal, because there are two forces acting in opposite directions. On the one hand, a major treatment advice becomes a stronger signal for being with a type b expert the more type g experts are in the market. Rejection would, therefore, be *more* attractive with more type g experts. This *revising effect* is especially strong when a major problem is not likely (low values of w). On the other hand, more type g experts make a rejection less worthwhile because the probability of being with such an expert becomes larger. According to this *direct effect*, a rejection would be *less* attractive with more type g experts. Summing up, we observe for high values of g that the mixed strategy equilibria become less likely with a higher percentage of type g experts. The reverse, however, holds true for low values of g : a higher value of g makes an equilibrium with mixed strategies more likely since the revising effect is then stronger than the direct effect.

This leads us to

[12] Notice that $\partial \left[\frac{wg}{(1-w)(1-g)\,x+w} \right] / \partial w > 0$.

Corollary 2: *A Bayesian-Nash equilibrium in mixed strategies is more likely*
(i) *the higher the price differential* $p_H - p_L$,
(ii) *the lower the switching costs* k, *and*
(iii) *the lower the probability* w *for an* H*-treatment.*
(iv) *For* $g > \frac{1-\sqrt{w}}{1-w}$, *a Bayesian-Nash equilibrium in mixed strategies is more likely, the lower the fraction of type* g *expert. The reverse holds true for* $g < \frac{1-\sqrt{w}}{1-w}$.

Since the equilibrium in mixed strategies exhibits a smaller amount of fraud than the fraudulent equilibrium in pure strategies, the following Corollary is tightly connected with Corollary 2. It delivers the relationship between parameters and the amount of fraud.

Corollary 3: *The amount of fraud in equilibrium decreases*
(i) *the higher price dispersion,*
(ii) *the lower switching costs* k,
(iii) *the lower probability* w *of an* H*-treatment, and*
(iv) *the lower the fraction* g *of the honest experts, as long as* $g > \frac{k+(p_H-p_L)}{2(p_H-p_L)}$. *The reverse holds true for* $g < \frac{k+(p_H-p_L)}{2(p_H-p_L)}$.

Note the different cut-off value of g in Corollaries 2 and 3. Again there exist two effects concerning the parameter g. First, a dishonest expert cheats more frequently, the more honest experts are in the market. Second, a rejection is favorable for the consumers, the more experts of type g are present, especially when the price differential is high. Therefore, cheating is less possible. The latter effect is responsible for the cut-off difference. The absolute value of the price differential $(p_H - p_L)$ becomes crucial for the sign of the derivative.

3 Efficiency versus equity

3.1 Welfare loss of inefficient equilibria

As listed in Table 1, there are two types of efficient equilibria for $k \neq 0$: both equilibria involve the consumer's pure strategy of always accepting the expert's advice in the first period. For $p_H = p_L$, the expert always tells the truth, and for $0 < p_H - p_L < \gamma\, k$, the type b expert always cheats. Accordingly, the first equilibrium is without fraud and the second one involves the maximum amount of fraud.

Despite their efficiency properties, both equilibria have a disadvantage. First, a price setting of $p_H = p_L$, which is needed in the non-fraudulent equilibrium, is very expensive for the consumers, especially for $c_H \gg 0$. Let us think of an operation as the major problem solving method versus an antibiotic treatment to solve the minor problem. The marginal cost differential is likely to be more than thousand Francs.[13] The consumers are, therefore, not eager to completely prevent the experts from cheating. It is cheaper for them to accept a certain amount of fraud in return for a lower price p_L, since the utility of the consumers is falling function of price p_L (see figure 2). Second, according to figure 3, the fraudulent equilibrium provides a much higher payoff for the type b expert than for the type g. This contradicts common ethics: doing a non-fraudulent, i.e., better, job should give at least the same payoff.

In order to lower the payoff of the type b expert, one has to assure that the mixed equilibrium is played. This mixed equilibrium, however, involves a welfare loss originating from the switching costs of the rejecting consumers. These efficiency costs of playing the mixed equilibrium are the total search costs s:

$$s = k\, y^*(w + (1 - w)(1 - g)\, x^*). \qquad (9)$$

Rejection is only chosen for a major advice, which arises when the customer has an H-problem or when an L-problem-customer meets a type b expert who cheats. The probability for this event is $(w + (1 - w)(1 - g)\, x^*)$. Given an H-advice, the customer rejects with probability y^* facing switching costs of k.

For given values of g, w, k, and c_H the problem for a regulator is to set the prices p_H and p_L in order to reach a desired outcome. To simplify the analysis, we set price p_H equal to c_H. There is no conflict with the equal compensation principle since the existence of a major problem is both observable and verifiable.[14] Accordingly,

[13] One may argue that an operation is observable and incurs higher marginal cost than c_L, however, if the problem is in fact minor, it is likely that the marginal costs are smaller than c_H, so that the markup is still larger for performing the major problem solution. In section 4, we show that observability does not change the results substantially. Nonobservability is more realistic in consulting services, for example.

[14] see Milgrom & Roberts (1992).

the experts are not tempted to perform the minor treatment for a customer with a major problem, although they only enjoy a positive mark-up with the minor treatment. Moreover, it is reasonable to set p_H as low as possible to minimize the incentive cheat.[15] Therefore, the search costs can be regarded as a function of p_L. In what follows, we assume that the fixed costs per expert are low enough in order to be compensated by the mark-up of the minor service p_L. That is, the mixed equilibrium can be sustained with a price differential of $p_H - p_L$ equal to $c_H - p_L$.

3.2 Minimizing the search costs

Playing the mixed strategies is appealing in terms of equity among experts, even though it is costly and could be avoided with an appropriate price setting. Furthermore, second opinions are often observed in reality. Given that the mixed equilibrium is played, it seems natural to minimize the welfare loss. The search costs are minimized for $p_L^* = p_H - \frac{2k}{g}$. The price p_L^*, given the other parameters g, w, k, and $p_H = c_H$, is, however, not always part of an equilibrium in mixed strategies. Equation (6) constrains the possible value of p_L. This constraint is binding for

$$w > \frac{1-g}{2-g} \tag{10}$$

If the constraint is binding, i.e., when the probability for the major problem is high relative to the fraction of the type g experts, the search costs are minimized for $\overline{p}_L$. Remember that $\overline{p}_L$ is the cut-off price where the kind of equilibrium changes. In this case, the experts choose $x^* = 1$ and the search costs amount to $\overline{s} = s(\overline{p}_L) = \gamma\,(1-g+wg)\,k^2/p_H$. The consumers are indifferent between playing the equilibrium in pure strategies and the one in mixed strategies. The type b experts, however, lose the amount of $\overline{L}_b = L_b(\overline{p}_L) = \gamma\,(1-w)\,gk$ due to the fact that the equilibrium with mixed strategies is played (see figures 2 and 3). Not only do they bear the whole efficiency loss $\overline{s}$, but they also give a positive amount to the type g experts. In the mixed equilibrium, a fraction of clients is changing from a type b expert to a type g expert.

[15] Notice that Wolinsky (1995) with endogenous price setting observes the only mark-up with the less expensive services as well.

These consumers provide additional mark-ups to the type g experts. Accordingly, type g experts are always better off when the equilibrium in mixed strategies at $\overline{p}_L$ is played.

If the constraint is not binding the consumers benefit from playing the mixed equilibrium. The type b experts cheat with probability $x^* = w/(1-w)/(1-g) < 1$. Furthermore, they fully bear the search costs in this equilibrium. The search costs amount to $s^* = s(p_L^*) = 4\,k^2 w/(g\,p_H)$. Additionally, the type b experts 'pay' a certain amount to the consumers as well as to the other type of expert. Accordingly, the population of the type b experts loses the amount of $L^* = 2k\,(1-2w)/g$ in total. The type g experts are, however, not always better off. Although they now receive the mark-up of the L-problem from former b-type clients, the mark-up itself becomes smaller. Due to the lower price p_L, the earnings of the g types may reduce.

- insert figure 4 about here -

Figure 4 displays the search costs as a function of p_L for two different parameter environments. Constraint (10) is binding for the minimum of s_1. Accordingly, $\overline{p}_1$ results as a solution to this optimization problem. The solution to the minimization problem of s_2 turns out to be p_2^*, because the constraint is not binding here.

3.3 Minimizing the cost of redistribution

In order to analyze the costs of this redistribution, it is useful to build the ratio $Q = (1-g)\,L/s$. Q measures the cost of a redistribution in terms of the welfare loss. That is, the mixed equilibrium redistributes the amount of $(1-g)\,L$ from the type b experts to the type g experts and to the customers. Moreover, this equilibrium exhibits a welfare loss of s originating from the search costs of the clients rejecting in the first period. The value of Q is larger than 1 since by playing the mixed equilibrium the type b experts are the only losers. The higher Q the more effective is the redistribution. Accordingly, the amount transferred from the type b experts to their colleagues and to the consumers is high relative to the efficiency loss s. The first order condition for maximizing Q with respect to p_L is strictly negative for all parameter values. Therefore, the most effective redistribution is achieved with the lowest possible price for the minor treatment $\underline{p_L}$.

This price must guarantee nonnegative profits for the experts, i.e., it should be possible to cover the fixed costs by selling the minor service.

In order to summarize these results we conclude with the following Corollary. It presupposes that efficient equilibria are unwanted in terms of equity or distribution. In order to balance the experts' earnings, the parameter environment must allow for the mixed equilibrium being played.

Corollary 4: *Two price setting strategies for a regulator to achieve a mixed equilibrium are considered:*
(i) set p_L such that the efficiency loss is minimized, and
(ii) set p_L such that the cost of the redistribution is minimized.
The solution to (i) is p_L^* or $\overline{p}_L$ depending whether equation (10) is binding or not. Problem (ii) requires that p_L is at its minimal feasible value $\underline{p_L}$.

4 Extension: Model with observable service

4.1 The basic model

The information problem between expert and client is reduced when the client is able to observe the expert's services. In this section, we analyze the expert-client relationship presupposing observability. That is, the expert cannot cheat by overcharging but he is able to cheat by overtreating. Although the information problem weakens to some extent, most of the model's features remain the same. In the following, we highlight the differences between the overcharging and the overtreating model.

Overtreatment is only profitable if the mark-up by performing the H-treatment is higher than the mark-up for the L-treatment. Therefore, a fraudulent equilibrium, i.e., an equilibrium with overtreatment, only exists for $p_H - c_H > p_L$.

We consider the mixed strategy equilibrium of this game. In the first period, the customers reject a major service with probability y in order to set a type b expert indifferent between overtreating and telling the truth when facing a client with a minor problem, i.e.,

$$(1 - y)(p_H - c_H) = p_L.$$

The profit of type b experts who tell the truth is certainly p_L. By cheating he may earn $p_H - c_H$. The client's rejecting strategy (y) determines the probability of earning the larger amount.

Accordingly, the customer randomizes over rejecting (y) and accepting ($1-y$) with

$$y^* = 1 - \frac{p_L}{p_H - c_H} \tag{11}$$

As in the model without observability, the type b experts randomize between cheating (x) and telling the truth ($1-x$) with

$$x^* = \frac{w\,k}{(1-w)\,(1-g)\,(g\,(p_H - p_L) - k)} \tag{12}$$

since the utility of the customers is independent of experts' marginal costs.

The two probabilities $y^* \in [0,1]$ and $x^* \in [0,1]$ constitute a Bayesian-Nash equilibrium when at least one probability is strictly higher than 0 or strictly lower than 1.

Proposition 3: *There is a Bayesian-Nash equilibrium in mixed strategies. The type b experts choose strategy $\widetilde{H}/L$ with probability x^* and the clients reject an H-advice with probability y^* for*

$$0 < \gamma\,k \equiv \frac{k\,(1-g+wg)}{g\,(1-g)(1-w)} < p_H - p_L, \tag{13}$$

provided that $c_H < p_H - p_L$, where $\gamma\,k$ is the cut-off price difference.

Summarizing the results at hand, Table 2 concludes this section presenting an overview of the different types of equilibria. Notice that the pure, fraudulent equilibrium (b) is not efficient anymore since type b experts generate superfluous costs c_H by overtreating.

type of equilibrium	strategies	parameter environment	behavior of type b expert	efficiency
a1	mixed	$0 < \gamma\,k < c_H < p_H - p_L$	fraudulent	inefficient
a2		$0 < c_H < \gamma\,k < p_H - p_L$		
b	pure	$0 < c_H < p_H - p_L < \gamma\,k$	fraudulent	inefficient
c	pure	$0 \leq p_H - p_L < c_H$	non-fraudulent	efficient
d	pure	$0 = k = p_L < c_H < p_H - p_L$	fraudulent	inefficient
e1	pure	$0 = k < p_L < c_H < p_H - p_L$	non-fraudulent	efficient
e2	pure	$0 = k < c_H < p_L < p_H - p_L$	non-fraudulent	efficient

Table 2: Properties of the different equilibria

4.2 Overcharging versus overtreatment

In our model, the difference between overcharging and overtreatment crucially depends on the observability of services. Overcharging is possible if the service is not observable. In contrast, overtreatment is the only possibility of cheating for observable services. It is, however, important to note that our overcharging model fits a broader concept of overtreatment, namely the searching for mark-ups. As long as fixed costs require strictly positive mark-ups for services, our overcharging model covers the main features of the overtreatment problem. Overtreatment in our model is narrowly defined as selling a major service when a minor service is needed. It is not profitable, however, when the mark-up of the two services is equal because only one service can be sold. In practice, more than one service can be billed and overtreatment cannot be avoided as easy as in our model. Therefore, the following difference between the two kinds of cheating should not suggest two mutually exclusive concepts.

As expected, the information problem between expert and client is more likely to be solved when the performed service is observable. Specifically, the non-fraudulent equilibrium becomes more likely in this case (see equilibrium c). The incentives for opportunistic behavior are lessened by the marginal costs of the major service. Accordingly, it is sufficient to set prices such that $p_H - c_H \leq p_L$ in order to prevent fraudulent behavior. This price setting does not reduce consumers' utility like in the nonobservable case which requires $p_H = p_L$ to provide proper incentives. Note that $p_L < p_H - c_H$ does not guarantee a non-fraudulent equilibrium in the nonobservable case since there is always an incentive to overcharge as long as the two prices are not equalized. The latter involves a high price for the minor service which must be paid for any service. This is the most important difference between the two cases.

Nevertheless, if experts have an incentive to overtreat, fraudulent equilibria arise for strictly positive switching costs. Furthermore, the outcome of the fraudulent equilibrium is not efficient anymore. Due to the experts' overtreatment, additional marginal costs arise, which are wasted. This makes the equilibrium (a) in mixed strategies more attractive than the equilibrium (b) in pure strategies compared to the model without observability. Not only are the equity properties of mixed equilibria more favorable, but also the efficiency comparison is not clear-cut anymore. There is a parameter environment for which the mixed equilibrium is more efficient than the pure equilibrium.

This is likely to happen for low values of the switching costs and for high values of the marginal costs of the major service.[16]

The profits of the experts and the utility of the consumers behave in similar ways as in the model without observability. Therefore, figures 1 to 4 also apply when the experts' services are observable. In figure 2, however, is now a range of large values of p_L that supports a non-fraudulent equilibrium. Accordingly, we obtain a jump in the consumer's utility up to the benchmark value for this price range as shown in figure 5.

- insert figure 5 about here -

In figure 4, the value of price p_L^* is more likely to be higher than $\overline{p}_L$ comparing to the case without observable services. The condition (10) binds stronger than the analogue in the observation case for most parameter environments. Consequently, an interior solution is more likely for observable services.

Contrary to the overcharging model, the consumers are always better off in the non-fraudulent equilibrium in case of observability. Here, the first best solution is also appealing in terms of equity. Therefore, we should discuss the cases for which the first best cannot be reached in spite of full observability. First, it is difficult to set prices such that the incentive to overtreat is not given anymore. Due to technological progress, the marginal costs of the treatment change over time whereas the prices are fixed in the short run. Second, our model strongly simplifies matters. We assume the treatment of just one problem. The incentive to overtreat persists if additional problems may be treated which provide positive mark-ups. This would be a case of genuine overtreatment. In our model, we only analyze overtreatment as solving a more severe problem than necessary.

And last, full observability is not very likely at all. Moreover, perfect observability is not even sufficient for avoiding overcharging. The performed service must be interpreted correctly. Otherwise, an expert is capable of selling a minor service as a major one. Hybrid cases in terms of observability are therefore plausible. Marginal costs may change when the problem to be solved is minor in fact and the observability is not

[16] The same is true for low values of w and g. The lower the probability of requiring the major service and the lower the fraction of the honest experts, the more attractive is the mixed equilibrium in terms of efficiency.

perfect. The treatment of a minor problem causes lower marginal costs as shortcuts are possible, e.g., precautionary measures can be omitted. To put it differently, the marginal costs of performing the major service are lower than c_H, because solving the minor problem is less difficult despite applying the same treatment. The expert's profit from cheating is then neither p_H, nor $p_H - c_H$ but something inbetween. As a result, the incentive for overtreatment persists.

To summarize, perfect observability does not alter the features of the equilibria. The plausibility of certain equilibria, however, is put in question in our setup with only two services.

5 Conclusion

In our setting, the expert client information problem can be solved in terms of efficiency and honest behavior under restrictive assumptions: First, all experts are honest when prices are set such that fraudulent behavior does not pay. Second, the clients are able to discipline the experts by rejecting every major treatment recommendation when the switching costs are zero, and the mark-up for the minor service is strictly positive. As long as one of these two assumptions does not hold true anymore, a certain amount of fraud cannot be avoided. Zero switching costs are not plausible. Changing the expert always entails certain costs. Moreover, equalizing prices is too expensive in the nonobservable case when the differential of the marginal costs between the two services is large. It is cheaper for the consumers to allow for some fraud in return for a lower price of the low-cost-service. Therefore, a fraudulent equilibrium results.

An efficient, fraudulent equilibrium has the disadvantage of an unequal income distribution between honest and dishonest experts. Furthermore, the amount of fraud is at its maximum value in this equilibrium. In contrast, the inefficient, fraudulent equilibrium is attractive because it balances the earnings of honest and dishonest experts. In addition, the resulting welfare loss is fully borne by the cheating experts. The inefficient, fraudulent equilibrium involves mixed strategies for the client as well as the cheating expert. To support this kind of equilibrium, we need a large enough price differential between the two services relative to the switching costs. A positive probability of consumers' rejection ensures that the experts of type b are cheating less frequently and that their earnings decrease. It is a interesting feature of this equilibrium that the inefficiency, which is actually caused by the cheating experts, is also borne by them.

We consider two price setting strategies for a regulator. Both price settings assure an equilibrium in mixed strategies to be played. Which price p_L should be chosen depends on the type b experts. If the type b experts are not regarded as 'bad guys' but are seen as just reacting to wrong economic incentives, it is reasonable to minimize the efficiency loss of the mixed equilibrium rather than the cost of the redistribution. Moreover, the type g experts also suffer from a low price p_L. Therefore, a price which minimizes the welfare loss may be more attractive for distribution reasons between consumers and experts. Otherwise, the cost of redistribution should be minimized. This aim is reached for the lowest feasible price of the minor service.

When the service is observable, prices should be set such that overtreatment does not pay. As long as overtreatment is profitable, the arguments about fraudulent equilibria hold true as in the nonobservable case with overcharging. The observable case is, however, not very interesting in our model. With observability, the first best solution is easy to attain. The incentive to cheat disappears for equalized mark-ups, because it is not possible to treat more than one problem. Moreover, the first best is also appealing in terms of equity since consumers' utility is highest in the first best. Nevertheless, as long as opportunistic behavior pays, our model without observability covers the main features of overcharging as well as overtreatment.[17]

The lesson we learn from this study is that the price differential between services of different costs should not be too small. Although the expert's incentive to cheat is higher for a large price discrepancy, the client's incentive to search for a second opinion is larger as well. We saw that a small fraction of 'bad' experts is able to cash in on consumers who accept anything. Second opinion collectors help balancing the earnings of fraudulent and non-fraudulent experts. Furthermore, the fraudulent experts cheat less frequently. Therefore, it is necessary to encourage the consumers to receive second opinions. This can also be achieved by lowering the switching costs. The latter is best done by offering a second opinion at a lower price. The switching costs may be non-pecuniary. First-opinion-experts who have to deliver all the information concerning the diagnosis[18], for example, lower the price of a second opinion substantially. In

[17]Cheating behaviour pays if observability is not perfect or if more than one service with strictly positive mark-up may be sold.

[18]We mean diagnosis in a broad sense. It could be any kind of information about the client which

practice, some health insurance companies provide free second opinion diagnoses for certain operations which makes sense in view of our analysis.

In addition, this study provides insights of a fair compensation schedule assuming heterogeneity among experts. A fair compensation schedule cannot be exploited by experts who deviate from the standard of the profession. That is, cheating experts should not be better paid than their honest colleagues. The mixed equilibrium we considered balances the experts' earnings across types. Accordingly, the presented price setting which sustains this mixed equilibrium constitutes a fair compensation schedule.

6 Appendix

Normal form of the game

minor problem with probability $(1-w)$:

Expert Customer	$\widetilde{H}/L$	$\widetilde{L}/L$
accept	p_H $B-p_H$	p_L $B-p_L$
reject	0 $B-gp_L-(1-g)p_H-k$	0 $B-gp_L-(1-g)p_H-k$

major problem with probability w:

Expert Customer	$\widetilde{H}/H$	$\widetilde{L}/H$
accept	p_H-c_H $B-p_H$	p_L-c_H $B-p_L$
reject	0 $B-p_H-k$	0 $B-p_L-k$

Expert has a weakly dominant strategy to choose $\widetilde{H}/H$ rather than $\widetilde{L}/H$, because he has to solve the major problem.

Reduced Form: the consumer decides to accept/reject a treatment after receiving an H -advice.

is needed for an efficient advice.

Expert Customer	$\widetilde{H}/H : \Pr\left[\widetilde{H}/H\right] = w$	$\widetilde{H}/L : \Pr\left[\widetilde{H}/L\right] = (1-g)(1-w)$
accept	$p_H - c_H$ $B - p_H$	$p_L - c_H$ $B - p_H$
reject	0 $B - p_H - k$	0 $B - gp_L - (1-g)p_H - k$

Proposition 1

We need two conditions to establish a Nash equilibrium in strictly mixed strategies:

(i) $0 < y^* < 1$:

$0 < \frac{p_H - p_L}{p_H} \Leftrightarrow p_H - p_L > 0 \wedge p_H > 0$

$1 > \frac{p_H - p_L}{p_H} \Leftrightarrow p_H > p_H - p_L \Leftrightarrow p_L > 0$

(ii) $0 < x^* < 1$:

$0 < \frac{w\,k}{(1-w)\,(1-g)\,(g\,(p_H - p_L) - k)} \Leftrightarrow g\,(p_H - p_L) > k$

$1 > \frac{w\,k}{(1-w)\,(1-g)\,(g\,(p_H - p_L) - k)} \Leftrightarrow (1-w)\,(1-g)\,(g\,(p_H - p_L) - k) > w\,k$

$\Leftrightarrow \frac{k\,(1-g+wg)}{g\,(1-g)(1-w)} < p_H - p_L$

It remains to show that $\frac{k\,(1-g+wg)}{g\,(1-g)(1-w)} < p_H - p_L \Rightarrow g\,(p_H - p_L) > k$

$\frac{k\,(1-g+wg)}{g\,(1-g)(1-w)} < p_H - p_L \Leftrightarrow k < \frac{g\,(1-g)(1-w)(p_H - p_L)}{(1-g+wg)}$

but $\frac{g\,(1-g)(1-w)}{(1-g+wg)} < g$ since $(1-g)(1-w) < (1-g+wg) \Leftrightarrow w > 0$.

Concluding, if $x^* < 1 \Rightarrow x^* > 0$

Proposition 2

ad (i) If $x^* \geq 1$, i.e., if the type b experts are always cheating, the customers always accept since there are no longer indifferent between accepting and rejecting. Experts are constantly cheating only when they are sure that it is not profitable for the clients to reject.

ad (ii) Experts have no incentive to cheat when prices are equalized. Accordingly, clients accept.

ad (iii) Consumers always reject an H -advice in the first period since it is for free ($k = 0$). Experts do not cheat first period customers, because it is more profitable to earn p_L than to earn nothing. All second period consumers have an H -problem and cannot be cheated.

ad (iv) When the lower service has no markup, the clients are not able to discipline experts even for $k = 0$, since the experts only make profits by selling the high price service.

Corollary 1-3

The only source of an efficiency loss are the switching costs k. For $k = 0$ all equilibria are efficient. For the pure strategy equilibrium in which the customers always accept, the switching costs are avoided.

(a) Comparative statics (Corollary 2)

(i) $0 < \gamma\, k \equiv \frac{k(1-g+wg)}{g(1-g)(1-w)} < p_H - p_L$

(ii) $\frac{\partial \gamma\, k}{\partial k} = \frac{1-g+wg}{g(1-g)(1-w)} > 0$, i.e., the mixed equilibrium becomes less likely as k rises.

(iii) $\frac{\partial \gamma\, k}{\partial w} = \frac{k}{g(1-g)(1-w)^2} > 0$, i.e., the mixed equilibrium becomes less likely as w rises.

(iv) $\frac{\partial \gamma\, k}{\partial g} = \frac{k(2g-g^2+wg^2-1)}{g^2(1-g)^2(1-w)} < 0$, because $1 - 2g + g^2 - wg > 0 \Leftrightarrow g < \frac{1-\sqrt{w}}{1-w}$, i.e., the mixed equilibrium becomes more likely as g rises when the latter condition holds. Otherwise $\frac{\partial \gamma\, k}{\partial g} > 0$.

(b) Comparative statics (Corollary 3)

(i) $\frac{\partial x^*}{\partial (p_H - p_L)} = \frac{-wgk}{(1-g)(1-w)(g(p_H-p_L)-k)^2} < 0$

(ii) $\frac{\partial x^*}{\partial k} = \frac{wg(p_H-p_L)}{(1-g)(1-w)(k-g(p_H-p_L))^2} > 0$,

(iii) $\frac{\partial x^*}{\partial w} = \frac{k}{(1-g)^2(1-w)(k-g(p_H-p_L))^2} > 0$,

(iv) $\frac{\partial x^*}{\partial g} = \frac{wk(2g(p_H-p_L)-k-(p_H-p_L))}{(1-g)^2(1-w)(k-g(p_H-p_L)^2} > 0 \Leftrightarrow 2g(p_H-p_L)-k-(p_H-p_L) > 0 \Leftrightarrow (p_H-p_L) > \frac{k}{2g-1}$ or $g < \frac{k+(p_H-p_L)}{2(p_H-p_L)}$

Proposition 3

In addition to Proposition 1, we need the condition $c_H < p_H - p_L$ in order to establish an equilibrium in strictly mixed strategies. For $c_H > p_H - p_L$, experts do not have an incentive to overtreat since the the price differential does not compensate the additional marginal costs.

Corollary 4

Minimizing the efficiency loss

minimize $s = k\,y^*(w + (1-w) * (1-g)\,x^*)$ or $s = \frac{kwg(p_H - p_L)^2}{p_H\,(g(p_H - p_L) - k)}$

FOC: $(p_H - p_L)(g(p_H - p_L) - 2k) = 0 \Leftrightarrow p_L = p_H, \frac{gp_H - 2k}{g}$

SOC: $\frac{2k^3wg}{p_H\,(g(p_H - p_L) - k)^3} > 0$

It remains to show that $g(p_H - p_L) - k > 0 \Leftrightarrow g(p_H - p_L) > k$

$\frac{g\,(1-g)(1-w)\,(p_H - p_L)}{1-g+wg} < g(p_H - p_L) \Rightarrow \exists$ only equilibria in mixed strategies: $k < g(p_H - p_L)$

$\frac{g\,(1-g)(1-w)\,(p_H - p_L)}{1-g+wg} < g(p_H - p_L) \Leftrightarrow \frac{(1-g)(1-w)}{1-g+wg} < 1 \Leftrightarrow (1-g)(1-w) < 1 - g + wg$

$1 - g - w + wg < 1 - g + wg \Leftrightarrow -w < 1 \forall w$

Constraint binding if $\overline{p}_L < p_L^*$, i.e.,

$\overline{p}_L = p_H - \frac{k\,(1-g+wg)}{g\,(1-g)(1-w)} < p_H - \frac{2k}{g} = p_L^*$

$\frac{1-g+wg}{(1-g)(1-w)} > 2 \Leftrightarrow 1 - g + wg > 2(1 - w - g + wg)$

$0 > 1 - g + wg - 2w \Leftrightarrow w > \frac{1-g}{2-g}$ □

search costs in equilibrium: $s^* = \frac{4k^2w}{g\,p_H}$.

Comparison of the welfare losses for the observability case

overtreatment loss (pure): $op = (1-w)(1-g)c_H$

overtreatment loss(mixed): $om = (1-w)(1-g)\left[(1-y^*)c_H + y^*(1-g)c_H\right]$

search loss (mixed): $s = ky^*(w + (1-w)(1-g)x^*$

efficiency loss (mixed): $em = om + s = \frac{wk\left[g(p_H - p_L)^2 - c_H((1-g)p_H - c_H + 2gp_L)\right]}{(g(p_H - p_L) - k)(p_H - c_H)}$

efficiency loss (pure): $ep = op + s = op$

$em < ep \Rightarrow k < \kappa = \frac{gc_H(1-w)(1-g)(p_H - p_L)(p_H - c_H)}{wg(p_H - p_L)^2 + c_H((1-g)(p_H - c_H) + wg(2p_L - p_H))}$

$\kappa > 0 \Leftrightarrow wg(p_H - p_L)^2 + c_H((1-g)(p_H - c_H) + wg(2p_L - p_H)) > 0$

case 1:$2p_L > p_H \Rightarrow \kappa > 0$

case 2:$2p_L < p_H \Rightarrow \kappa > 0$, iff $p_H > 1$ (sufficient)

Conclusion: there is a parameter environment where the efficiency loss of the mixed equilibrium is smaller than the efficiency loss of the pure equilibrium.

References

[1] DARBY, M. R. and KARNI, E., Free Competition and the Optimal Amount of Fraud, *Journal of Law and Economics* **16** (1973), 67-88.

[2] EMONS, W., Credence Goods and Fraudulent Experts, *Rand Journal of Economics* **28** (1997), 107-119.

[3] JAFFEE, D.M., and RUSSELL, T., Imperfect Information, Uncertainty, and Credit Rationing, *Quarterly Journal of Economics* **90** (1976), 651-666.

[4] MARTY, F. E., Capacity as a Determinant of the Supply for Physicians' Services, *Diskussionsschriften Universität Bern* **98-5** (1998).

[5] MILGROM, P. and ROBERTS, J. Economics, Organization & Management, New Jersey: Prentice Hall, (1992).

[6] PESENDORFER, W., and WOLINSKY, A., Second Opinions and Price Competition: Inefficiency in the Market for Expert Advice, Mimeo (1998).

[7] PITCHIK, C. and SCHOTTER, A., Honesty in a Model of Strategic Information Transmission, *American Economic Review* **77** (1987), 1032-1036.

[8] WOLINSKY, A., Competition in a Market for Informed Experts' Services, *Rand Journal of Economics* **24** (1993), 380-398.

[9] WOLINSKY, A., Competition in Markets for Credence Goods, *Journal of Institutional and Theoretical Economics* **151** (1995), 117-131.

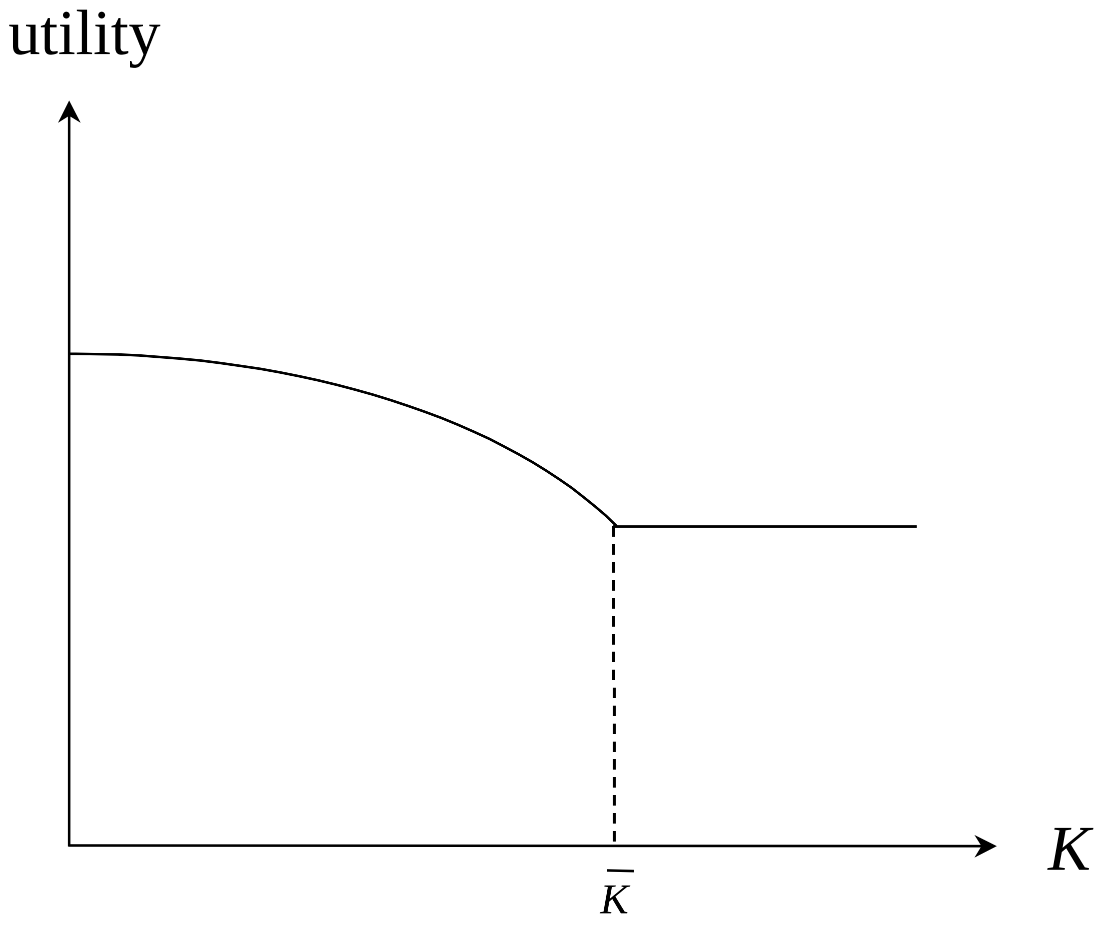

Figure 1: Utility as a function of K

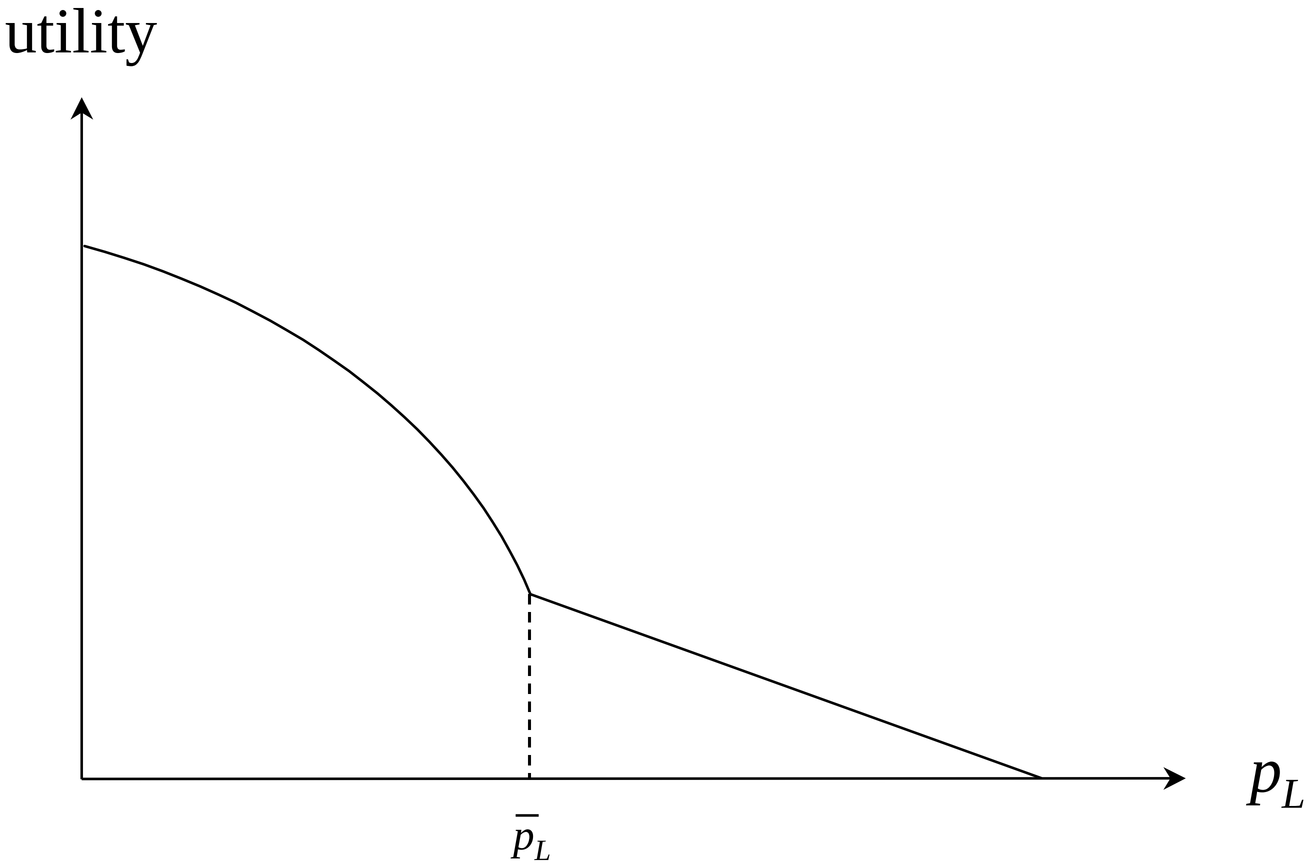

Figure 2: Utility as a function of p_L

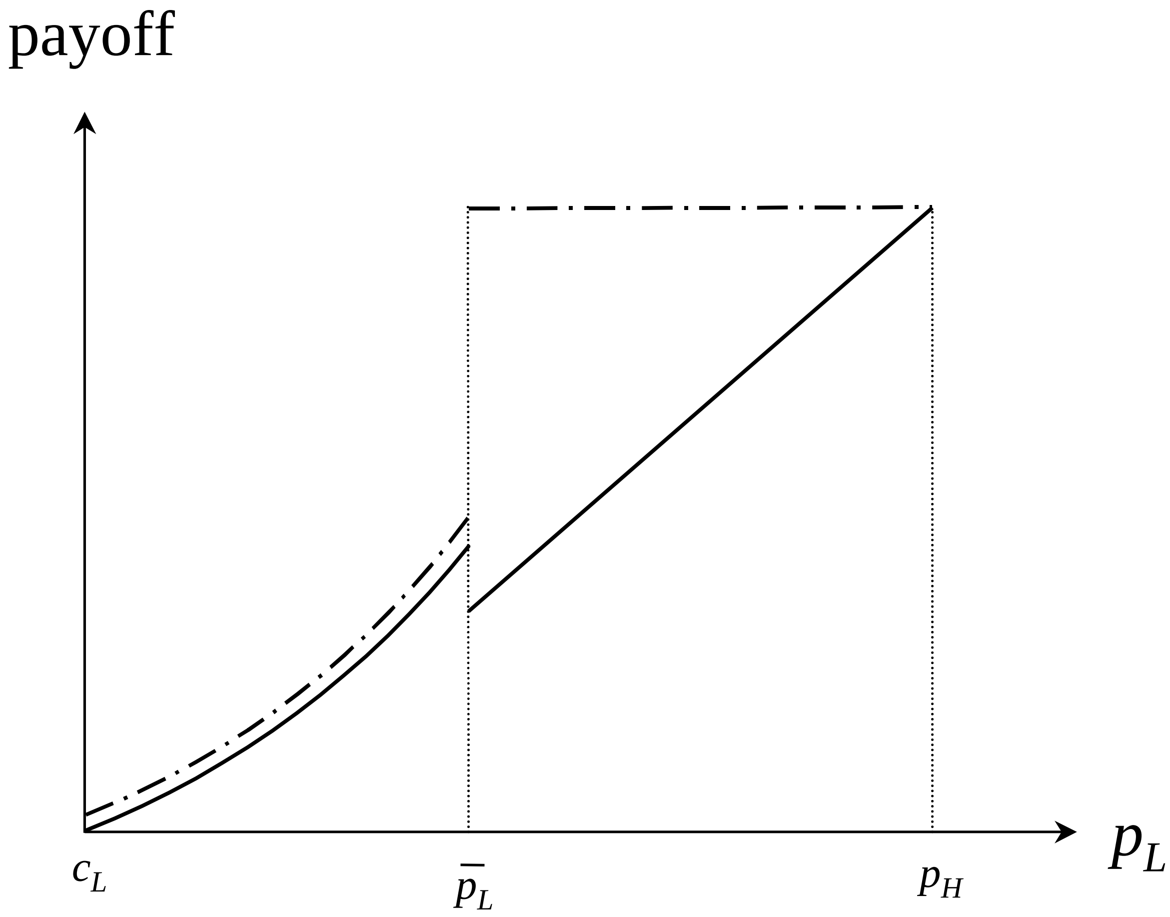

Figure 3: Population payoff of *g*-type experts ——

Population payoff of *b*-type experts — · — · —

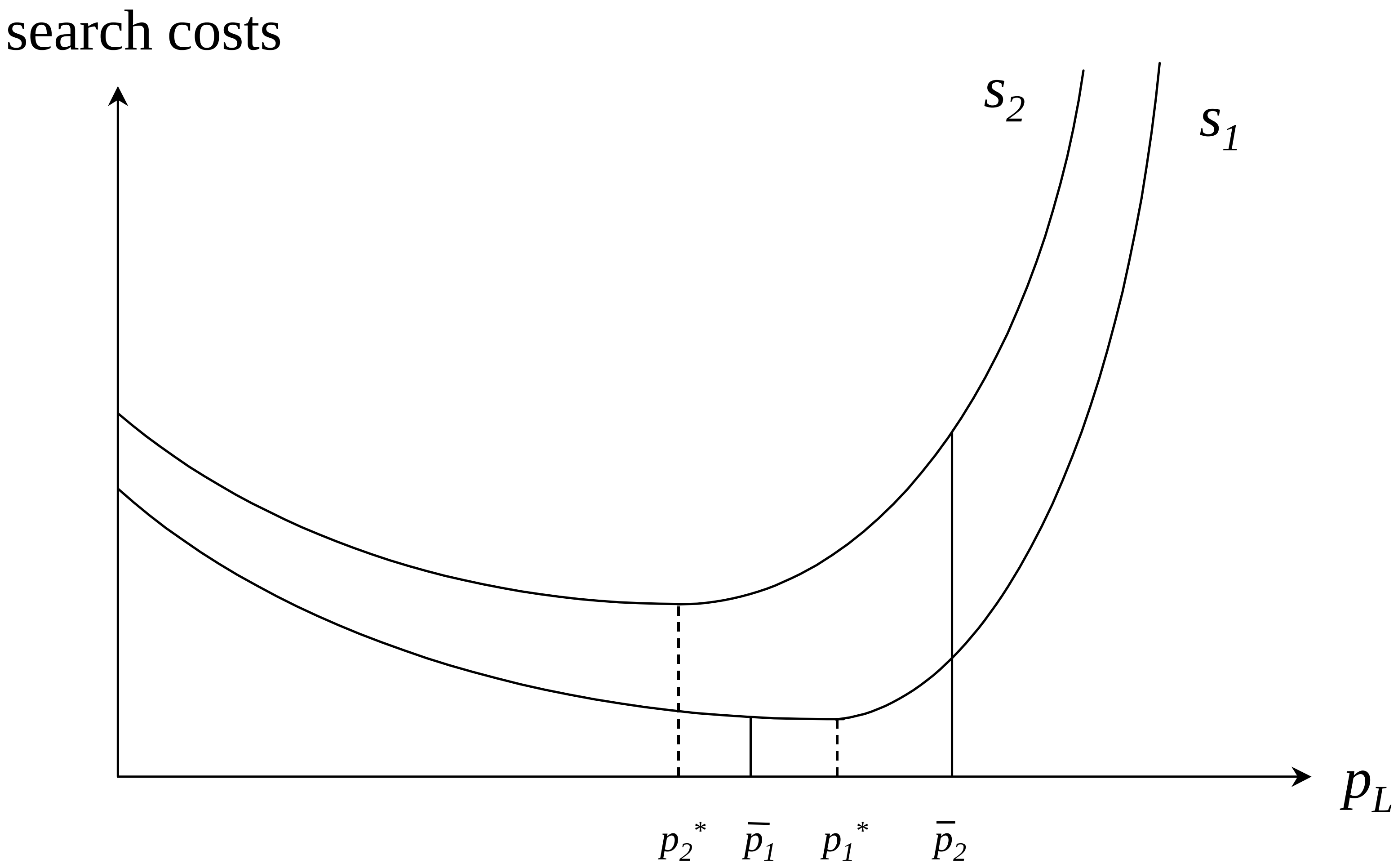

Figure 4: Minimizing search costs

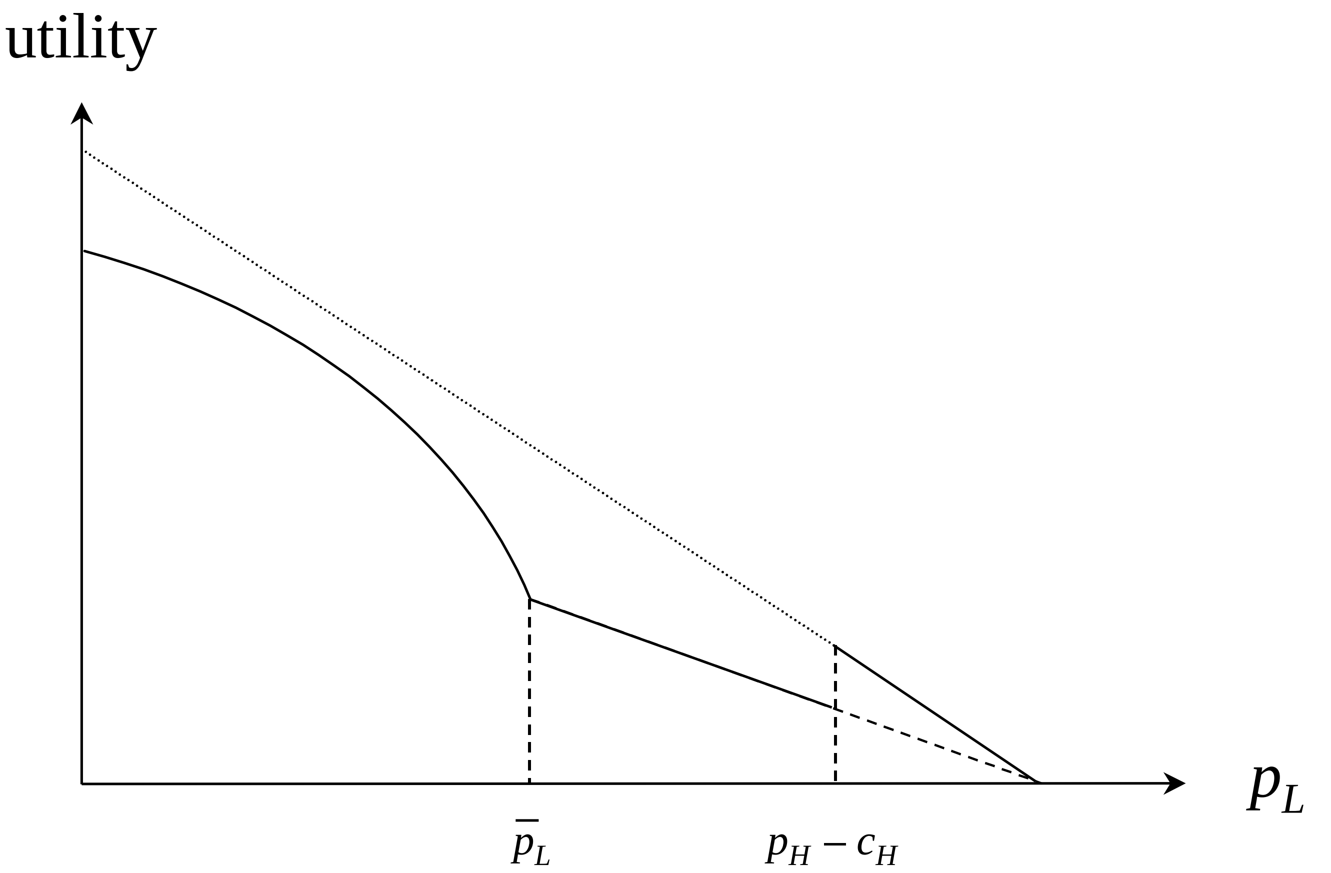

Figure 5: Utility as a function of p_L (observable case)