© *The Journal of Risk and Insurance*, 2005, Vol. 72, No. 1, 159-176

Insurance in a Market for Credence Goods

Kai Sülzle
Achim Wambach

Abstract

We study the impact of variations in the degree of insurance on the amount of fraud in a physician-patient relationship. In a market for credence goods, where prices are regulated by an authority, physicians act as experts. Due to their informational advantage, physicians have an incentive to cheat by pretending to perform inappropriately high treatment levels leading to overcharging patients. Our approach aims on the impact on changes in each, patients' and physicians' incentive structure when the proportional degree of coinsurance varies. It is shown that a higher coinsurance rate may lead to either less fraud in the market and a lower probability of patients searching for second opinions or more fraud and more searches.

Introduction

This article deals with the provision of expert services in a physician-patient framework. Physicians act as experts when patients do not know the exact medical services they need, and physicians therefore determine how much and which type of service will be demanded. The literature labels such services or goods as credence goods (e.g., Darby and Karni, 1973). This feature can also be observed in markets for legal and repair services where customers *ex post* often cannot determine if they were served appropriately or not. A particular consequence in this situation with asymmetric information is supplier-induced demand, which is beside moral hazard, one of the main informational problems in the market for health services. It refers to situations where either physicians can and do treat more than what is medically necessary or charge for a more expensive treatment than the one they actually provide, as they have better information available than the patients. Moral hazard describes the case where patients demand more than necessary, due to the fact that they are insured and thus pay only a small fraction of the bill. In both cases it is argued that a higher degree of coinsurance weakens the problems created. If patients have to pay more, they consume less, which reduces their moral hazard. If patients have to pay more

Kai Sülzle is at the Ifo Institute for Economic Research, Poschinger Str. 5, 816679 München, Germany. Achim Wambach is at the Department of Economics, University of Erlangen-Nuremberg, P.O. Box 3931, 90020 Nürnberg, Germany. The authors can be contacted via e-mail: suelzle@ifo.de and achim.wambach@wiso.uni-erlangen.de. We thank the participants of the ESEM 2003 conference in Stockholm for their helpful comments.

they also have a higher incentive to control their physician, which in turn reduces the potential for supplier-induced demand. We investigate this latter effect and show that this conclusion does not hold in general.

We study the impact of insurance arrangements on the degree of fraud in such a market. Our exclusive aim lies on the changes in patients' and physicians' behavior, since we are dealing with fraud within the patient-physician relationship. In particular, we ask whether a higher coinsurance rate on the side of the patients makes fraud in this market more or less likely. In our model, physicians diagnose patients. As patients do not have any information about the degree of illness, physicians can claim that the illness is very severe even if only a small treatment is necessary. The choice variable of the patients is to accept or reject the diagnosis, in which case they go to another physician. As a first intuition, one would suggest that an increase in the coinsurance rate will lead to more patients searching for a second opinion if they face a diagnosis for a large treatment. The reasoning would be that a higher coinsurance rate gives patients a larger financial incentive to search, as they have to pay a larger fraction of the medical bill. More patients searching for second opinions give physicians an incentive to diagnose more honestly. However, this intuition is only partially correct.

In particular, we show that: If the coinsurance rate increases, patients are *ceteris paribus* more willing to reject a high diagnosis and to search for a second opinion, which is in agreement with the intuition just given. However, in equilibrium, there are two possible consequences of an increase in the coinsurance rate: Either patients search less and physicians diagnose more honestly or patients search more and physicians diagnose less honestly.

Underlying to these results is that physicians have to take two considerations into account when deciding whether to diagnose honestly or not. First, the reaction of patients. If patients are less willing to accept a high diagnosis, this makes fraudulent behavior less attractive. Second, the behavior of other physicians. If other physicians behave relatively dishonest, and patients often reject their first diagnosis, the chance is high that a patient coming to a physician is already on his second visit. In this case, the patient would accept a high diagnosis as a confirmation of the first diagnosis. This in turn makes fraudulent behavior more attractive.

In equilibrium, the extent of fraud has to be just sufficient to keep patients indifferent. The interaction of the two effects described above then can go in both directions: Either all other physicians diagnose more honestly in which case each individual physician is also better off diagnosing more honestly, and patients thus have less incentives to search for a second opinion. Or, all other physicians recommend a higher treatment more frequently, and patients search more often, which increases the chance that a patient is on his second visit when coming to a physician. This in turn increases the incentive to recommend the high treatment.

We discuss four further aspects in this framework. First, if physicians can specialize, then an increase in the coinsurance rate enhances the possibility to make specialization sustainable. The reason is that specialization is a possible response to the informational asymmetry in this market (see Wolinksy, 1993). If a physician is not able to do special treatments, she will only refer the patient if the required treatment is truly severe. Now, if the coinsurance rate is larger, the problem of dishonest diagnosis is more severe from the patient's point of view, thus favoring specialization. Second, we study a complete insurance contract where it is shown that if insurance can condition both on the charged treatment by the physician and on the number of physicians visited, then fraud can be eliminated. Third, we modify the behavioral assumptions made by allowing for some honest physicians who always make the correct treatment recommendations. It is shown that with honest physicians around, opportunistic physicians might have even stronger incentives to make a false treatment recommendation, as patients are more willing to accept it. Finally, while our main analysis assumes risk-neutral patients, we also investigate risk aversion. It can be shown that the main results on the equilibrium behavior remain to hold. However, there is an interesting new result on the partial equilibrium behavior. Risk-averse patients might *ceteris paribus* be less willing to search for a second opinion if the coinsurance rate increases. The reason is that a second opinion comes with an income risk, which is larger if the coinsurance rate is larger.

Related Literature

Models on Credence Goods. Several articles deal with credence goods in markets, for example medical services (see, e.g., De Jaegher and Jegers, 2001; Dulleck and Kerschbamer, 2001; Emons, 1997, 2001; Pesendorfer and Wolinsky, 1999; Wolinsky, 1993, 1995). However, to our knowledge the only contributions considering the possibility of insurance in such a context are Dionne (1980, 1984).[1] Although Dionne studies the reaction of patients to changes in the insurance structure, he does not analyze how exactly physicians modify their behavior and its resulting effect on patients' behavior.

Models on Insurance Fraud. Insurance fraud is a well-studied research topic. Most dominant in this line of research are costly state verification models (see, e.g., Townsend, 1979; Mookherjee and Png, 1989; Fagart and Picard, 1999) and costly state falsification models (see, e.g., Lacker and Weinberg, 1989; Crocker and Morgan, 1998). In the former case, the insurer can validate claims only at a cost; while in the latter case, the insured can manipulate claims at a cost to him. While much attention has been given to the incentive of the insured to betray the insurer, only little has been done to investigate in how far third parties have an incentive to manipulate a claim (see, e.g., Brundin and Salanié, 1997). In this literature, it is assumed that the third party colludes with the insured. Apart from the above-mentioned articles by Dionne no work has been done on insurance fraud, where the third party has an informational advantage compared to the insured.

The remainder of this article is structured as follows: The assumptions of the model are presented in section "The Model," where we also determine patients' and physicians' optimization problems as well as the equilibrium outcome. Our main results on the impact of insurance on the degree of fraud are derived. And the "Discussion" section concludes the article.

[1] De Jaegher and Jegers (2001) comment shortly upon copayments by the patients. They conclude that higher out-of-pocket costs lead to less fraud.

THE MODEL

The framework we use is a simplified version of the model by Wolinsky (1993) where we additionally introduce insurance.

There is a continuous number of patients in the market, indexed by $s \in [0, 1]$. Conditional of being ill, a patient suffers a big medical problem with probability $p \in (0, 1)$ and a small problem with probability $(1 - p)$. Although a patient notices when having a problem he cannot specify its gravity.

There are $I \geq 2$ identical physicians in the market who can offer diagnosis and treatment. The physicians' intention is to maximize the profit per patient. When treating a patient with a big problem, a physician incurs costs H whereas the costs for treating a small problem are L with $L < H$. In the remainder, we label the treatment of a big problem H and that of a small problem L.

We assume that physicians charge exogenously fixed prices H for the H treatment and $L + e$ for the minor problem L, with e standing for an additional charge (*mark-up*) on the L treatment that ensures that there is an incentive to provide the L treatment honestly. The reason for the fixed price assumption is that in existing markets for medical services we can indeed observe in many countries fixed prices arising from a bargaining process on a central level between lobbyists (e.g., U.S. Medicare organizations) and government or from legal regulations (e.g., as in Germany). Furthermore, we want to focus on insurance and fraud but not on price setting. Concentrating on fixed prices ensures the existence of at least one equilibrium with a positive level of fraud (see Wolinsky, 1993; Pitchik and Schotter, 1987).[2]

The interaction between patients and physicians proceeds as follows:

1. A patient s who is ill chooses a physician. If he faces an H diagnosis, he accepts this with probability $y(s) \in [0, 1]$. With probability $(1 - y(s))$ he visits another physician for a second opinion which he accepts with certainty.[3] Visiting a physician, a patient has to incur search and waiting costs k.
2. Physicians diagnose those patients visiting and give a treatment recommendation. Their recommendation policy is described by $x_i \in [0, 1]$, that is with probability x_i a patient having an L problem receives a recommendation for an H treatment from physician i. A patient with an H problem always obtains an H recommendation. If patients accept, they get treatment with prices according to the diagnosis. Otherwise, they leave to get a second opinion.

Payoffs are as follows: A patient who is treated obtains utility $B - \tau P - nk$, where B is the benefit of treatment, P is the price, $\tau \in (0, 1)$ is the coinsurance rate, and $n \in \{1, 2\}$ is the number of physicians the patient visits. We assume that B is large enough such that any patient will undergo treatment eventually. Note that k encompasses all

[2] With flexible prices, equilibria with fraud are much harder to sustain (see, e.g., Dulleck and Kerschbamer, 2001).

[3] The model could be generalized to allow rejections of second opinions also (see, Wolinsky, 1993).

the costs the patient has to bear fully. This might be search and waiting costs, but nonmonetary costs of undergoing a diagnosis like pain.[4]

A physician who treats a patient with diagnosis L makes a profit of e. If the diagnosis is H and the patient indeed suffers an H illness, his profit is zero. Finally, if the reported diagnosis is H, but the patient only suffers a minor L illness, then the physicians profit is $H - L$, which is per assumption higher than e. In this case, the patient gets defrauded by the physician who pretends to perform an H treatment while only applying an L treatment. Accordingly, since the patient cannot observe the exact treatment he gets overcharged.

Note that the model presented here—like the model in Wolinsky (1993)—describes overchargement, i.e., patients have to pay more than necessary without obtaining the more expensive treatment. With minor modifications the model can also be applied to the case of overtreatment, where patients obtain a H treatment although a L treatment would have been sufficient. This could be seen, if one considers a margin on the H treatment (say e^*), which is larger than the margin on the L treatment ($e^* > e$). In such a case, a physician could make a higher profit from applying and charging an inappropriate H treatment to an L patient instead of applying the adequate L treatment. Most of the analysis presented below would go through in such a model (but see also footnote 8).

As a benchmark case, consider the situation with full and no insurance. In the case of full insurance ($\tau = 0$) there would not be an incentive for patients to ever go to a second physician, since treatment costs would be fully borne by the insurance and a patient who visits a second physician would only face further search and waiting costs k. Physicians would defraud all L patients, since those would stay for sure at their first visit. The case of no-insurance ($\tau = 1$) is equivalent to the case in Wolinsky (1993) and is included in the following analysis by setting $\tau = 1$.

We now turn to the case with a positive coinsurance rate. We concentrate on symmetric equilibria where all physicians choose the same recommendation policy x and all patients have the same acceptance rate y.

Patients' Optimization Problem

A patient's expected utility is maximized when expected costs C_P are minimized through optimal choice of acceptance policy y. C_P is composed as follows:

$$
\begin{aligned}
C_P &= k + (1-p)(1-x)\tau(L+e) + [p + (1-p)x]y\tau H \\
&\quad + [p + (1-p)x](1-y)\left[k + \frac{(1-p)x(1-x)\tau(L+e) + [p + (1-p)x^2]\tau H}{p + (1-p)x}\right]. \quad (1)
\end{aligned}
$$

The search costs k have to be paid for the first visit to a physician in any case. With probability $(1-p)(1-x)$ a patient gets a correct L diagnosis—which is accepted with certainty—and incurs treatment costs $\tau(L+e)$. With probability $[p + (1-p)x]$ a patient

[4] The model could be extended to allow for monetary costs of diagnosis that is partially borne by the insurance. This would not modify the general results of this analysis.

gets an H diagnosis and accepts with probability y from the first physician leading to a payment of τH. The remaining part of (1) describes the situation when a patient gets an H recommendation from the first physician—which is the case with probability $[p + (1-p)x]$—but declines with probability $(1-y)$ and consults a second physician, which leads to costs k again. The following fraction weights the payments for an L and an H treatment by the second physician with their conditional probabilities given that the first physician diagnosed an H problem. From Equation (1) it follows that a patient will accept an H diagnosis with probability 1 (0) whenever

$$\tau H < (>) k + \frac{(1-p)x(1-x)\tau(L+e) + [p + (1-p)x^2]\tau H}{p + (1-p)x}. \tag{2}$$

When both sides are equal, the patient is indifferent between accepting or rejecting. When accepting the patient incurs τH, when declining he faces expected costs described on the right-hand side of (2).

Lemma 1: *Consider a symmetric recommendation policy $x \in [0, 1]$. If k is sufficiently small, the patients' best response correspondence is*

$$y^*(x) \in \begin{cases} \{1\} & \text{if } x \in [0, x_1) \cup (x_2, 1], \\ [0, 1] & \text{if } x \in \{x_1, x_2\}, \\ \{0\} & \text{if } x \in (x_1, x_2), \end{cases}$$

with

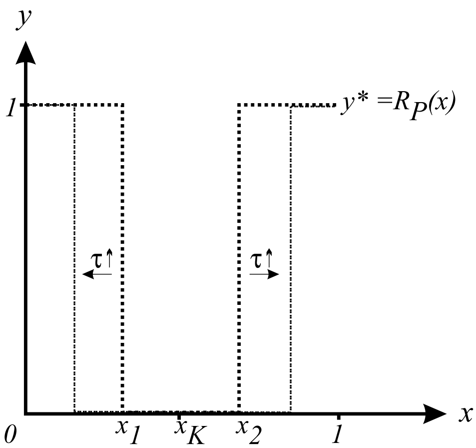
$$x_{1,2} = \frac{1}{2}\left(1 - \frac{k}{(H-L-e)\tau}\right) \pm \sqrt{\frac{1}{4}\left(1 - \frac{k}{(H-L-e)\tau}\right)^2 - \frac{p}{1-p}\frac{k}{(H-L-e)\tau}}. \tag{3}$$

The proof is relegated to the Appendix.

Lemma 1 shows that if physicians diagnose relatively honestly (i.e., x is small) patients will accept the high diagnosis with probability 1, as it is then quite likely that he is indeed an H patient. Interestingly, if physicians cheat very often (i.e., x close to 1), patients will also accept the high diagnosis. The reason is that searching for a second opinion does not help very much—also the second physician is expected to lie about the true diagnosis. Looking for a second opinion is a good strategy only for intermediate values of x. In this case the patient has a good chance to be treated honestly by the second physician while the chance that the first physician lied is not negligible.

In the Appendix, we give a sufficient and necessary condition for Lemma 1 to hold. From Equation (3) it can be seen that a necessary condition is that $k \leq \tau\,(H - L - e)$. If search costs k were larger than the right-hand side, which is the maximum potential financial gain of visiting a second physician ($\tau\,(H - L - e)$), then the agent would always accept any diagnosis. Consequently, with large k (or small τ, respectively) the only equilibrium is where all physicians always give an H diagnosis, i.e., $x = 1$, and the patient always accepts (i.e., $y = 1$). In the following, we concentrate on the case where k is small enough (or τ is large enough).

FIGURE 1
Patients' Reaction Correspondence



In Figure 1, we plotted the patients' reaction correspondence.

We now come to our first proposition concerning the impact of insurance on the behavior of the market participants.

Proposition 1: *Ceteris paribus, the higher the coinsurance rate τ the higher is a patient's incentive to decline an H diagnosis from the first consulted physician and to choose $y = 0$.*

Proof: Setting the equal sign in expression (2) and taking the partial derivative of x with respect to τ gives:

$$\frac{\delta x}{\delta \tau} = -\frac{(H - L - e)(1 - x)x}{(1 - 2x)(H - L - e)\tau - k}. \tag{4}$$

The numerator of this derivative is positive. The whole fraction is positive (negative) if $(1 - 2x)\,(H - L - e)\tau - k > (<)\; 0$, where the critical value is given by

$$x_K = \frac{1}{2}\left(1 - \frac{k}{(H - L - e)\tau}\right). \tag{5}$$

Comparing (5) with the two solutions $x_{1,2}$ in Lemma 1 shows that $x_1 < x_K < x_2$ leading to $\frac{\delta x_1}{\delta \tau} < 0$ and $\frac{\delta x_2}{\delta \tau} > 0$. Hence an increase in τ indicates a change in patients' reaction correspondence as shown in Figure 1. The values x_1 and x_2 for which a patient is indifferent are closer to the extrema 0 and 1—that is there is a broader range of x-values between x_1 and x_2 for which the patients' best response is $y = 0$ as stated in Proposition 1. Q.E.D.

Proposition 1 is intuitively clear—if the patient has more to gain from obtaining a second opinion, he is *ceteris paribus* more likely to decline an H recommendation. In Figure 1, we indicated that an increase in τ shifts the two points where the patient is indifferent to the outside. Intuitively one would suggest that such an increase in search for second opinions triggers a decrease in false recommendation, as physicians should fear more patients leaving. However, as we will see in the next subsections, in a full equilibrium analysis this argument does not necessarily hold.

Physicians' Optimization Problem

We consider the situation of a physician who wants to maximize her expected profit. Due to the fixed price structure the physician can only make a profit with L patients. The choice variable of physician j is the individual recommendation policy for L patients, which is denoted by x_j. Her optimization problem is then given by maximizing the following profit per L patient π_j:

$$\pi_j = (1 - x_j)e + x_j \left(\frac{y + x(1-y)}{1 + x(1-y)} \right) (H - L), \tag{6}$$

where x stands for the average market level of fraud. With a correct diagnosis the profit is the mark-up e while the expected profit of an incorrect H diagnosis is $(H - L)$ with probability $[y + x\ (1 - y)]/[1 + x\ (1 - y)]$. The fraction $\frac{1}{1+x(1-y)}$ of all patients are on their first visit when consulting a specific physician, hence they accept an H diagnosis with probability y. The remaining $\frac{x(1-y)}{1+x(1-y)}$ patients are on their second (and last) visit to a physician and therefore accept either diagnosis with certainty.

The physician will choose $x_j = 1$ ($x_j = 0$) whenever

$$e < (>) \left(\frac{y + x(1-y)}{1 + x(1-y)} \right) (H - L). \tag{7}$$

She is indifferent between a honest diagnosis of an L patient and an incorrect recommendation of an H treatment when both sides of expression (7) are equal. With a correct L diagnosis a patient stays at the first physician with certainty, which leads to a net profit of e. When receiving an H diagnosis the patient only stays with probability $\frac{y+x(1-y)}{1+x(1-y)}$ which results in the payoff $(H - L)$ for the physician.

Lemma 2: *If $e < \frac{H-L}{2-y}$, the symmetric physicians' best response correspondence is given by*

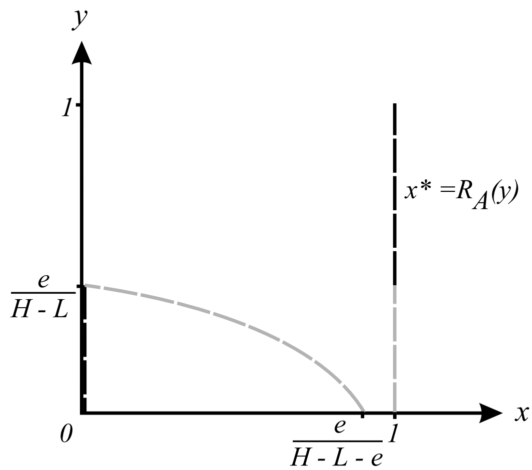
$$x^*(y) \in \begin{cases} \{1\} & \text{if } y \in \left[\frac{e}{H-L}, 1 \right], \\ \left\{ 0, \frac{e - y(H-L)}{(1-y)(H-L-e)}, 1 \right\} & \text{if } y \in \left[0, \frac{e}{H-L} \right], \end{cases}$$

if $e \geq \frac{H-L}{2-y}$ the symmetric physicians' best response correspondence is given by

$$x^*(y) \in \begin{cases} \{1\} & \text{if } y \in \left[\frac{e}{H-L}, 1 \right], \\ \{0\} & \text{if } y \in \left[0, \frac{e}{H-L} \right]. \end{cases}$$

The proof is relegated to the Appendix.

FIGURE 2
Physicians' Equilibrium Strategy



If the patients accept an H recommendation sufficiently often (i.e., y is large), then physicians prefer to cheat ($x = 1$). If on the other hand, patients tend to go for a second opinion (i.e., y is close to zero), then physicians prefer to diagnose honestly ($x = 0$). If e, the markup on the L treatment, is small, then there also exists a region where all physicians are indifferent between cheating or not.

Figure 2 gives a graphical interpretation of Lemma 2. The black dotted line shows those symmetric best response strategies that always exist. The gray line is only part of the symmetric best response strategies if $e < \frac{H-L}{2-y}$.

In this article, we want to investigate the effect of insurance on the level of fraud. Therefore, the interesting case for us is when $e < \frac{H-L}{2-y}$. Only then equilibria where physicians do false recommendations with a positive probability smaller than 1 are possible.

Equilibrium Discussion

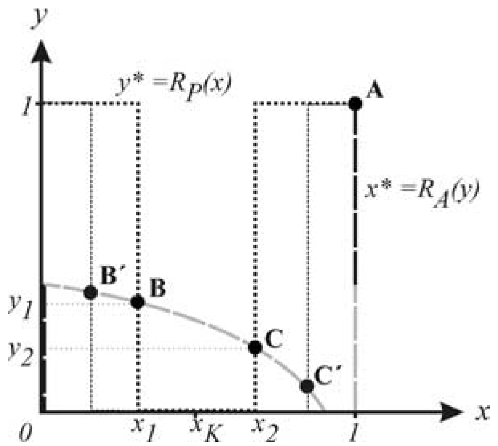
Joining the results from the previous two subsections gives us a situation as illustrated in Figure 3. We can describe the equilibrium outcomes as follows.

Lemma 3: *In the described game there exists always an equilibrium in pure strategies where physicians defraud every L patient by diagnosing an H problem* ($x = 1$). *Patients always accept an H diagnosis on their first visit* ($y = 1$).

Additionally for a sufficiently small k and $e < \frac{H-L}{2-y}$ *there exist two equilibria in mixed strategies where physicians defraud L patients with positive probability and patients accept an H diagnosis with positive probability from the first physician.*

Proof: We get three equilibria as can be seen in Figure 3. Equilibrium A is in pure strategies: Patients always accept an H diagnosis from the first physician consulted. Physicians always recommend an H diagnosis to every L patient.

FIGURE 3
Equilibria in a Market With Fraud



The two additional equilibria are in mixed strategies. In B, the probability of fraudulent advice is relatively low at x_1 while in C the probability is relatively high at x_2. In both cases, the patient is just indifferent between accepting an H diagnosis and searching for a second opinion.[5] Note that depending on the size of $\frac{e}{H-L-e}$ the intersection point C can also lie on the x axis between x_1 and x_2. This would indicate an equilibrium where physicians randomize between a correct L diagnosis and defrauding whereas patients always decline an H diagnosis on their first visit. Q.E.D.

Although in outcome B physicians diagnose more honestly than in C, it is not obvious in which equilibrium total fraud is minimized, as patients are also more willing to accept wrong recommendations in equilibrium outcome B. Formally, the number of fraudulently wrong diagnoses is given by $xy + x(1-y)x$. The first term describes those patients who are fraudulently diagnosed at their first visit (with probability x) and accept this diagnosis (with probability y). The second term is those who reject the first diagnosis (with probability $1-y$) but are wrongly diagnosed also at their second visit (with probability x), which they then accept. Using the equilibrium condition for $x^*(y)$ from Lemma 2 for an equilibrium in mixed strategies this expression transforms to $x(1-y)\frac{e}{H-L-e}$. From this it follows that in outcome B the total amount of fraud is minimized, as both x and $1-y$ are lower than in outcome C.

We now turn to comparative statics of the analysis. The impact of an increase in the coinsurance rate on the outcome in the market is shown in our main proposition.

Proposition 2: *An increase in the coinsurance rate τ leads to the following changes in the three possible equilibrium outcomes:*

[5] The same equilibrium outcome could be attained in an asymmetric equilibrium in pure strategies where a proportion y_1 (or y_2) of patients always accept an H recommendation, while a proportion $(1-y_1)$ (or $(1-y_2)$) declines any recommendation. Similarly for the physicians, where a proportion x_1 (or x_2) always recommend an H treatment while $(1-x_1)$ (or $(1-x_2)$) recommend honestly.

1. *At outcome B′, there is less fraud in the market than in B. Additionally, more patients accept an H diagnosis from the first physician they consult (see Figure 3).*
2. *At outcome C′, there is more fraud in the market and less patients accept an H diagnosis from the first physician than in C. Outcome A is still an equilibrium outcome.*

Proof: The intuition of Proposition 2 can be immediately observed in Figure 3. As demonstrated in Proposition 1 an increase in τ leads to a direct change of patients' reaction correspondence—again illustrated by the thin line in Figure 3. There is no direct influence of a variation of τ on physicians' symmetric best response correspondence, but there is an indirect reaction due to the implicit change of patients' behavior. The physicians' indirect reaction to a variation of τ is given by[6]

$$\frac{\delta y}{\delta \tau} = -\frac{(H-L-e)^2(1-x)(1-y)x}{[(1-2x)(H-L-e)\tau - k][(H-L-e)x-(H-L)]}. \tag{8}$$

The numerator is positive whereas the whole fraction is positive or negative depending on the denominator. Since $(H-L-e)x-(H-L)$ is negative, we get the same critical value x_K as in (5); that is $\frac{\delta y}{\delta \tau} > 0$ if $x < x_K$ and $\frac{\delta y}{\delta \tau} < 0$ if $x > x_K$. Hence we distinguish two situations:

1. $x > x_K$: If there is already a relatively high level of fraud in the market, say x_2, the denominator is negative in $\frac{\delta x}{\delta \tau}$ and positive in $\frac{\delta y}{\delta \tau}$. Therefore, $\frac{\delta x}{\delta \tau} > 0$ and $\frac{\delta y}{\delta \tau} < 0$. An increase in τ leads to an increase in x and more L patients get defrauded as before. In addition, less patients accept an H diagnosis from the first diagnosing physician.
2. $x < x_K$: If the existing level of fraud in the market is relatively low, say x_1, the denominator is positive in $\frac{\delta x}{\delta \tau}$ and negative in $\frac{\delta y}{\delta \tau}$, leading to $\frac{\delta x}{\delta \tau} < 0$ and $\frac{\delta y}{\delta \tau} > 0$. An increase in τ leads to a decrease in x and less L patients get defrauded as before. In addition, more patients accept an H recommendation from the first physician. Q.E.D.

Intuitively one would presume that an increase in the coinsurance rate will lead to a higher willingness of patients to search for a second opinion. This in turn makes it less attractive for physicians to provide a false diagnosis. That is y and x will both decrease. This does not hold in equilibrium.

Starting from a situation where physicians are just indifferent between diagnosing honestly and fraudulently, a decrease in y makes it strictly better to behave honestly (i.e., to set $x = 0$). But for $x = 0$ it is a patients' best response to stay with their first physician (i.e., $y = 1$). One way to achieve the balance of these effects is shown in Figure 3 at points B and B′. Here, the market level of fraud is relatively low and few customers seek second opinions. Patients on their first visit do not reject expensive recommendations all the time but only with positive probability, because the current diagnosis is likely to be correct. Experts do not recommend expensive treatments to customers with minor problems all the time because only few second-time visitors

[6] This is obtained by totally differentiating expressions (2) and (7) and solving for $\frac{\delta y}{\delta \tau}$.

populate the market. However, there exists a second possibility to balance these effects as it is shown in Figure 3 at points C and C′. Here, the market level of fraud is relatively high and many customers seek second opinions. First-time patients do not reject expensive recommendations all the time because the next physician is likely to overcharge too, while experts do not overcharge all the time because first-time customers seek second opinions sufficiently vigorously.

An increase in the coinsurance rate gives patients *ceteris paribus* a larger financial incentive to reject expensive treatments, as they have to pay a larger fraction of the bill. So in equilibrium B′ physicians have to become even more honest to keep patients indifferent, while in equilibrium C′ they have to become even less honest, compared to equilibria B and C, respectively.

DISCUSSION

In this discussion, we generalize the assumptions made so far as follows: First, we discuss specialization of physicians and more complete insurance contracts as possible institutional responses to dishonest behavior. Second, further behavioral aspects are considered by introducing honest physicians and risk-averse patients.

As Wolinsky (1993) shows, one institutional response to dishonest behavior could be the specialization of physicians. If some physicians specialize on the L treatment, a patient visiting such a physician can be assured that if he obtains an H recommendation, this diagnosis is correct. The physician has no incentive to give a false recommendation, as she does not profit from it. A patient without insurance who visits a physician who can only provide L treatments must only pay for the low treatment in case his illness is not severe. However, he faces the risk of being referred to a specialist in which case he has to pay the search costs k once more. Wolinsky (1993) shows that in this framework specialization can be an equilibrium whenever the expected payment by the patient is lower than $H + k$, which would be the bill of a patient who directly goes to someone who can provide both treatments and charges H in any case.

Now insurance is introduced into this model. If a patient visits a physician who can only do the low treatment, then he still faces the risk of being referred to a specialist in which case he has to bear the full search costs by himself. On the other hand, if the patient would visit the specialist directly, the resulting higher payment is less severe, as he is partially insured against this. Thus we can conclude that the range for possible specialization equilibria increases with the coinsurance rate. If patients have to pay more by themselves they are more critical concerning the price of treatment, while for small values of τ it is only the search costs which matters for them.

A second response to fraudulent behavior could be a more complete insurance contract where the insurer conditions the contract on additional information. So far, the degree of coinsurance was the only choice variable. However, at least in the framework of this model, the insurer could also condition the insurance contract on the charged treatment and on the number of physicians visited.

In such a setting, there exists an equilibrium without fraud. The fraud eliminating contract takes the following form: In case that the charged treatment is L, the patient will be paid $L + e$ if only one physician was consulted, and $L + e + k$ if two physicians were consulted. On the other hand, if the charged treatment is H, then the insurer will pay less than H if only one physician was consulted, and $H + k$ if a second physician made this recommendation. With such an insurance contract, the patient will accept any L recommendation, since searching for a second opinion will not make him better off. However, if the patient is faced with an H recommendation from his first physician, he is better off searching for a second opinion, in which case all costs including search costs will be covered by the insurance, while staying with the first physician will cost him H and less than H is reimbursed.[7] Now, if all physicians treat honestly, then a physician who treats an L patient can be sure that this patient is on his first visit and will thus search for a second opinion if he is faced with an H recommendation. Thus honest treatment recommendation is a best reply. This insurance contract is designed to make the agent look for a second opinion whenever the treatment recommendation is high. The resulting outcome is the same as in the specialization case. Patients with an L illness are treated correctly by the first physician, patients with an H illness will visit a second physician.

We now generalize the behavioral assumptions made. Following Marty (1999), consider a setting with two types of physicians, where one type always acts honestly, while the other type defrauds patients with positive probability. It is obvious that if the number of honest physicians is small enough, then in analogy with outcomes A, B and C, still all three types of equilibria will exist. As can be seen from Figure 3, this is the case as long as the percentage of honest physicians is less than $1 - x_2$. As a second case, consider the situation where the percentage of honest physicians is larger than $1 - x_1$. As can be seen from Figure 3, in this case the patient will accept any treatment recommendation. Therefore, all profit maximizing physicians will make the wrong recommendation. This would be the unique outcome. Finally, if the percentage of honest physicians lies between $1 - x_2$ and $1 - x_1$, then profit maximizing physicians will mix between honest and fraudulent recommendations. The equilibrium outcome will then be the analogue to outcome B. All honest physicians give a correct diagnosis, profit maximizing physicians make a false diagnosis with positive probability, and patients are just indifferent between accepting and not accepting the recommendation.[8]

Let us further extend the analysis by introducing risk aversion, which after all is one of the reasons for the existence of insurance. As the reaction correspondence of the physicians is not modified by risk aversion, the equilibrium result, i.e., x decreases and y increases or the other way around, does not change. The intersection points will either move to the left or to the right. However, in contrast to the risk-neutral case a risk-averse agent might have *ceteris paribus* less incentives to search for a second opinion if the coinsurance rate increases. The intuition behind that claim is as follows:

[7] Note that the patient is fully insured, as he would never chose a treatment that will not be fully reimbursed.

[8] Another modification on the behavioral assumptions could be to allow for some informed patients. This raises new problems, however. In particular, it is quite unclear what happens if an informed patient is diagnosed wrongly. Does he leave, does he tell the physician, does he inform a regulatory authority? In case he informs other patients of this fraudulent behavior, one would have to consider reputation of physicians as well, possibly along the line of the reputation scenario in Wolinsky (1993). Telling the authority with an appropriate resulting punishment might be enough to induce the physicians to behave honestly.

a patient who rejects a recommendation and searches for a second opinion faces an income risk. He does not know whether the next diagnosis is L in which case his payment is τL or H, in which case he has to pay τH. If τ increases, this income risk increases. Although the expected income gain of search increases, the increase in risk might make it unattractive to search further.[9] Therefore, if a patient is risk averse it might be the case that *ceteris paribus*, the higher the coinsurance rate τ the lower the patient's incentive to decline an H diagnosis from the first-consulted physician. Hence, the opposite of Proposition 1 might hold under risk aversion.[10]

APPENDIX

Proof of Lemma 1: Take $x \in [0, 1]$ as given, i.e., an L patient receives a correct diagnosis with probability $(1 - x)$ and an incorrect H diagnosis with probability x. Rearranging (2) (with an equal sign instead of the inequality) and substitution with $a = \frac{k}{(H-L-e)\tau} < 1$ yields

$$x^2 - (1-a)x + \left(\frac{p}{1-p}\right)a = 0. \tag{A1}$$

The x_1 and x_2 defined in Lemma 1 are real solutions to Equation (A1) and will exist if

$$(1-a)^2 - 4\left(\frac{p}{1-p}\right)a > 0, \tag{A2}$$

which always holds for a small enough. It follows from expression (A1) that for both $x < x_1$ and $x > x_2$ a patient strictly prefers to stay with the first physician and chooses $y = 1$. For $x \in (x_1, x_2)$ it is better to leave the first physician by setting $y = 0$ when receiving an H diagnosis. Q.E.D.

Proof of Lemma 2: Depending on the patients' symmetric strategy y we distinguish three situations:

(a) $y = 1$: Patients play the pure strategy and always accept an H diagnosis from the first physician. Setting $y = 1$ in (7) and rearranging yields

$$-e + H - L > 0. \tag{A3}$$

Obviously it is always a physicians' best response to set $x = 1$ meaning to recommend every patient an H treatment.

(b) $y = 0$: Now patients always decline an H diagnosis from the first physician. But it is not possible for a physician to distinguish between patients coming for the first and second time. Rearranging (7) yields

[9] Here it is assumed that the premium the patient pays does not depend on the coinsurance rate. An increase in the coinsurance rate is most likely to be accompanied by a decrease in premium, which also alters risk aversion. We ignore this effect here.

[10] See the Appendix for a formal proof of this claim.

$$-e + \frac{x}{1+x}(H - L) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0, \tag{A4}$$

where x is the recommendation policy of the other physicians. We consider three cases:

1. When $x = 0$ all other physicians always recommend honestly an L treatment to an L patient what solves (A4) for

$$-e < 0. \tag{A5}$$

   Hence for the individual physician j it is also optimal to choose $x_j = 0$. When $x = 0$ no L patient receives an H diagnosis from an other physician, which implies that when an L patient visits physician j it is his first visit to a physician. If physician j would recommend an H treatment it would be declined with certainty and physician j would make no profit.
2. When $x = 1$ all other physicians always defraud L customers by recommending an H treatment. $x_j = 1$ is a best response when

$$\begin{aligned} -e + \frac{1}{2}(H - L) &\geq 0, \\ e &\leq \frac{1}{2}(H - L). \end{aligned} \tag{A6}$$

   When (A6) holds physician j sets $x_j = 1$ and also defrauds every L patient of her clientele. This is because now there are sufficiently many L patients in the market who are on their second visit and therefore accept an H diagnosis with certainty. When $e > \frac{1}{2}(H - L)$, then $x = 1$ cannot be a candidate for a symmetrical physicians' best response to $y = 0$, since unilateral deviation to $x_j = 0$ would lead to a higher payoff for physician j.
3. When $x \in (0, 1)$ all other physicians defraud their patients with an L problem with positive probability. For a symmetric best response, physician j must also be indifferent between diagnosing honestly and giving a false recommendation. From (7) this holds if

$$x = \frac{e}{H - L - e}. \tag{A7}$$

   A solution exists if $e < \frac{1}{2}(H - L)$. If $e > \frac{1}{2}(H - L)$ then no $x \in (0, 1)$ is a candidate for a best response to $y = 0$ due to the same argumentation as above.

(c) $y \in (0, 1)$: Patients mix between accepting and declining an H diagnosis from the first physician. Rearranging (7) (with an equal sign) yields

$$x(1 - y)(H - L - e) - e + y(H - L) = 0. \tag{A8}$$

From partial derivation we get

$$\frac{dx}{dy} = -\frac{(H-L) - x(H-L-e)}{(1-y)(H-L-e)} < 0. \tag{A9}$$

An increase in y leads to a decrease in x. With a higher y, a physician needs less L patients on their second visit to be indifferent between honest behavior and defrauding L patients.

For the best responses x to $y \in (0, 1)$ we again distinguish three cases:

1. Suppose all other physicians choose $x = 0$. There exists a y so that (A8) holds with $x = 0$:

$$y = \frac{e}{H-L}. \tag{A10}$$

   Therefore, for low values of y ($y \in (0, \frac{e}{H-L}]$), $x_j = 0$ is a best response as long as all other physicians choose $x = 0$.
2. Suppose all other physicians choose $x = 1$. From (A8) we get that $x_j = 1$ is a best response whenever it holds that

$$(1-y)(H-L-e) - e + y(H-L) = H - L - e - (1-y)e > 0. \tag{A11}$$

   This holds as long as $e < \frac{H-L}{2-y}$.
3. Suppose all other physicians randomize with some $x \in [0, 1)$. From (A8) we get indifference for physician j to diagnose honestly or not if it holds:

$$x = \frac{e - y(H-L)}{(1-y)(H-L-e)}. \tag{A12}$$

   The so determined x lies between zero and one as long as $e < \frac{H-L}{2-y}$. Because of (A9) the x in (A12) is decreasing in y and reaches $x = 0$ for $y = \frac{e}{H-L}$. Q.E.D.

Impact of Insurance in Case of Risk Aversion

To concentrate on the effects of insurance on patient behavior, simplify notation by writing $\alpha(1-\alpha)$ as the probability of receiving an L (H) diagnosis from a second physician. Now consider a patient who received an H recommendation from the first physician he visited. The utility of the patient of being treated is given by $U(B - \tau P - nk)$ where U is a concave function. The other parameters are as before.

For the patient to be indifferent between accepting the first diagnosis (H) and searching for a second opinion it must hold:

$$U(B - \tau H - k) = \alpha U(B - \tau(L+e) - 2k) + (1-\alpha)U(B - \tau H - 2k). \tag{A13}$$

In that case the utility from accepting the H treatment from the first physician (on the left-hand side) equals the expected utility of declining and searching for a second opinion (on the right-hand side). We will show that for given values of α an increase in τ can make accepting more attractive if the patient is risk averse.

The derivative of (A13) with respect to τ yields

$$-U'(B-\tau H-k)H \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} -\alpha U'(B-\tau(L+e)-2k)(L+e)$$
$$-(1-\alpha)U'(B-\tau H-2k)H. \tag{A14}$$

For a risk-neutral patient with constant marginal utility we get

$$-H < -\alpha(L+e)-(1-\alpha)H, \tag{A15}$$

so as previously shown a risk-neutral patient prefers to get a second opinion if τ increases.

Now consider a risk-averse patient. If the utility function displays decreasing absolute risk aversion the following inequality holds

$$-\frac{U'''}{U''} > -\frac{U''}{U'}. \tag{A16}$$

Applying Pratt's theorem (Pratt, 1964) it follows that a person with utility function $-U'$ is more risk averse than someone with utility function U. A comparison with Equation (A13) then gives

$$-U'(B-\tau H-k) > -\alpha U'(B-\tau(L+e)-2k)-(1-\alpha)U'(B-\tau H-2k), \tag{A17}$$

since a more risk-averse person prefers to get the safe bet whenever a less risk-averse person is indifferent between a safe and a risky bet. Now, if $(L+e)/H$ is sufficiently close to one, then also expression (A14) might hold with a larger sign, showing that a risk-averse patient might be less willing to search for a second opinion if the coinsurance rate increases, which shows that under risk aversion the opposite of Proposition 1 might hold. The proof is given by an example: Take $U(x) = \ln(x)$ and set $B = 100$, $H = 90$, $L = 40$, $e = 10$, $k = 5$, and $\tau = 0.8$. α is determined by (A13), which yields $\alpha = 0.24$. In this case, expression (A14) reads $-3.913 > -4.040$ which proves the claim. Q.E.D.

REFERENCES

Brundin, I., and F. Salanié, 1997, Fraud in the Insurance Industry: An Organizational Approach, Mimeo, Department of Economics, University of Toulouse.

Crocker, K., and J. Morgan, 1998, Is Honesty the Best Policy? Curtailing Insurance Fraud Through Optimal Incentive Contracts, *Journal of Political Economy*, 106: 355-375.

Darby, M. K., and E. Karni, 1973, Free Competition and the Optimal Amount of Fraud, *Journal of Law and Economics*, 16: 67-88.

De Jaegher, K., and M. Jegers, 2001, The Physician-Patient Relationship as a Game of Strategic Information Transmission, *Health Economics*, 10: 651-668.

Dionne, G., 1980, Analyse des effets de l'assurance et de la relation de confiance consomateur-producteur sur les possibilités d'abus des chirugiens, *L'Actualité Économique*, 56: 211-238.

Dionne, G., 1984, The Effects of Insurance on the Possibilities of Fraud, *The Geneva Papers on Risk and Insurance*, 9: 304-321.

Dulleck, U., and R. Kerschbamer, 2001, On Doctors, Mechanics and Computer Specialists or Where Are the Problems With Credence Goods, University of Vienna, Economics Working Paper No. 0101.

Emons, W., 1997, Credence Goods and Fraudulent Experts, *RAND Journal of Economics*, 28: 107-119.

Emons, W., 2001, Credence Goods Monopolists, *International Journal of Industrial Organization*, 19: 375-389.

Fagart, M. C., and P. Picard, 1999, Optimal Insurance Under Random Auditing, *The Geneva Papers on Risk and Insurance Theory*, 24: 29-54.

Lacker, J. M., and J. A. Weinberg, 1989, Optimal Contracts Under Costly State Falsification, *Journal of Political Economy*, 97: 1345-1363.

Marty, F., 1999, The Expert-Client Information Problem, Mimeo, Department of Economics, University of Bern.

Mookherjee, D., and I. Png, 1989, Optimal Auditing, Insurance and Redistribution, *The Quarterly Journal of Economics*, 104: 399-415.

Pesendorfer, W., and A. Wollinsky, 1999, Second Opinions and Price Competition: Inefficiency in the Market for Expert Advice, Mimeo, Department of Economics, Northwestern University.

Pitchik, C., and A. Schotter, 1987, Honesty in a Model of Strategic Information Transmission, *American Economic Review*, 77: 1032-1036.

Pratt, J. W., 1964, Risk Aversion in the Small and in the Large, *Econometrica*, 32: 122-136.

Townsend, R. M., 1979, Optimal Contracts and Competitive Markets With Costly State Verification, *Journal of Economic Theory*, 21: 265-293.

Wolinsky, A., 1993, Competition in a Market for Informed Experts' Services, *RAND Journal of Economics*, 24: 380-398.

Wolinsky, A., 1995, Competition in a Market for Credence Goods, *Journal of Institutional and Theoretical Economics*, 151: 117-131.