Schmollers Jahrbuch 124 (2004), 233 – 256
Duncker & Humblot, Berlin

Asymmetric Information and the Demand for Health Care – the Case of Double Moral Hazard

By Udo Schneider*

Abstract

The production of health does not only depend on the medical services supplied by the physician but is also influenced by the patient's compliance. A model of medical treatment is presented in which both the actions of physician and patient are modeled as a productive input. The analysis distinguishes between three cases of strategic interaction. The consequences of asymmetric information between physician and patient are lower activity levels, only in the case of strategic substitutes the result might change. Furthermore, the effects of the implementation of a demand-side coinsurance are discussed.

Zusammenfassung

Die Gesundheitsproduktion hängt nicht alleine von den in Anspruch genommenen medizinischen Leistungen ab, sondern wird auch durch die Compliance des Patienten beeinflusst. Das hier präsentierte Behandlungsmodell betrachtet sowohl die medizinische Leistung des Arztes als auch die Handlungen des Patienten als produktiven Input. In der Analyse wird zwischen drei möglichen Ausprägungen der strategischen Interaktion unterschieden. Beidseitige asymmetrische Information zwischen Arzt und Patient resultiert dann in einem Absinken des Niveaus beider Inputs, lediglich im Fall der strategischen Substitute kann es zu einem anderen Ergebnis kommen. Darüber hinaus wird die Einführung einer Selbstbeteiligung auf Patientenseite untersucht.

JEL-Classification: I 11, D 82

1. Introduction

The key relationship in a health care system is the relation between patient and physician. In many economic studies dealing with the health care sector

---

* For helpful comments I would like to thank Brit Albers, Walter Ried, Volker Ulrich, and an anonymous referee.

the physician has an informational advantage in supplying medical services.[1] A theoretical model applied to the physician-patient relationship is the principal-agent theory that deals with situations of asymmetric information and the delegation of tasks if comparative advantages exist.[2] Compared to the organization of health care systems in practice, the theoretical results for physician payment systems are rather complex (cf. Holmström [1979], Arrow [1985] and Zweifel et al. [2001]). Furthermore, the application of a standard principal-agent model to the health care system is somewhat problematic (cf. Schneider [2002]). In fact, one major difference between theory and practice is that in the health care sector several "complementary agents" (Zweifel [1994]) affect the physician-patient relationship. Among these agents, especially the insurance company plays an important role. There exist contractual arrangements that incorporate both key participants of the health care sector, physicians and patients (cf. Kortendieck [1993], Gaynor [1994], Börsch-Supan [1998] and Cutler / Zeckhauser [2000]). It follows that the insurance company performs the task of a mediator with the result that individual actions like the demand for medical care, the consumption of health care goods and its financing fall apart (cf. Wille / Ulrich [1991]).

Many models deal with the physician-patient relationship and incorporate an insurance company, e.g. Selden (1990), Blomqvist (1991), Ellis / McGuire (1990) and Ma / McGuire (1997). None of these models faces the problem of mutual asymmetric information between the actions of physician and patient. In detail, this means that the latter cannot evaluate the quality of the physician's services and the physician does not possess exact information about the treatment-accompanying behavior of the patient.[3] This health-related effort of the patient is called *compliance* that is chosen in addition to the medical services (cf. Wille / Ulrich [1991], 27). It is possible to characterize the relation between medical services and compliance by the concepts of *strategic substitutes, strategic complements* or *strategic independence* (cf. Bulow et al. [1985]). In the first case of strategic substitutes, an increase in the activity of one agent leads to a decrease in the marginal productivity of the other agent's activity. In contrast to this, actions are strategic complements if the marginal productivity increases due to an increase in the level of the other one's action. In the case of strategic independence, there is no effect on the marginal productivity of the other agent's action.

---

1 For literature concerning the physician's behavior see Dionne / Contandriopoulos (1985) and Pauly (1980).

2 Another point of view is that market failure due to asymmetric information is no reason for public intervention. Instead the failure should be interpreted as scarcity and handled by a market process (cf. Shmanske [1996], 197 ff.).

3 The patient acts as a consumer and a production factor. The treatment result is a "joint product" of medical services and individual utilization (cf. Wille / Ulrich [1991], 27).

The basic approach of this paper is to illustrate the relationship between the physician's medical services and the patient's compliance. Therefore, only the interaction in the case of a consultation is analyzed. The results have important implications for health policy because governmental interventions regulating the patient's health related behavior affect the use of medical services and vice versa.

The paper is organized as follows: the next section describes the theoretical basis and the model structure focusing on the equilibrium of the cooperative and the non-cooperative solutions. The third section deals with the relation between the patient's compliance and the medical services and looks at the implications of a demand-side coinsurance.

2. A model of treatment decisions

2.1 Basic structures

The model presented here is based on a paper by Cooper/Ross (1985) about product warranties and the care of buyers and sellers. In contrast to their presentation, risk-aversion is assumed for the patient and an insurance company is incorporated. The latter difference is important because Cooper and Ross analyze direct contracts between buyer and seller. In a health care system, no direct contractual relationship between physician and patient exists. Consumers and insurance companies close an insurance contract, whereas insurer and physician place a contract about the remuneration of medical services. First, conventional insurance contracts without a coinsurance are analyzed. Afterwards, an extension of the model compares these results to a situation with an implementation of a coinsurance for the patient.

The following model neglects the patient's decision about the consultation of a physician. Instead, the analysis concentrates on the treatment stage. After the patient's decision to choose medical treatment, both the physician and the patient pick their actions simultaneously. The patient's expected utility depends on his net income. This corresponds to his gross income (y) minus the insurance premium (σ) and his co-payment (βm), denoting the product of the coinsurance parameter $(0 < \beta < 1)$ and the medical services in monetary units. The gross income is positively related to the state of health G after medical treatment.[4] Consequently, the patient is able to obtain a high-income level in a good state of health whereas a bad state of health reduces his productivity and his income will decrease.[5] The insurance premium has to be

[4] It is assumed that the gross income is sufficiently large so that the net income is always positive.

[5] An example to think of is the payment of sickness benefits. If the patient cannot work because of an illness he will receive a transfer payment from the insurance com-

paid in all states of the world but the co-payment only in situations when the patient decides to visit the physician in order to receive treatment.

As a simplification, we consider two conditions for the patient's post-treatment health status: either the patient is healthy after treatment or he is still sick. In the first case the variable G takes a high value G_1 and in the second case a low value G_0, with $G_1 > G_0$. A good state of health is realized with probability $p \in (0, 1)$, a bad one with probability $(1 - p)$. That is why recovery is not a deterministic but a stochastic process. The probability of a good state of health can be written as $p(a, m)$, given the patient's compliance (a) and the medical services supplied by the physician (m). For that reason, $p(\cdot)$ can be interpreted as a health care production function. With regard to the two inputs (a, m) it is assumed that both reveal a positive but diminishing marginal productivity, i.e. through an increase of one input the probability of a good state of health rises, but at a decreasing rate. The important result of this formulation of the health care production function is that with higher inputs bad states of health become less likely but cannot be ruled out (cf. Schneider [2002]).

The patient's expected-utility function is additive-separable in utility resulting from net income $U(\cdot)$ and disutility of 'providing' the compliance $D(a)$, i.e. the more a patient supports the treatment the higher is the resulting disutility.[6] The patient is risk-avers in his net income and the corresponding utility-function is concave. The disutility is assumed a convex function of the compliance. Expected utility can be written as:

$$(2.1)\quad EU = p(a,m)U\big(y[G_1] - \sigma - \beta m\big) + \big(1 - p(a,m)\big)U\big(y[G_0] - \sigma - \beta m\big) - D(a)\ .$$

Subsequently, U_1 is defined as utility resulting from a good state of health (G_1) and U_0 as utility from a bad state of health (G_0):

$$U_0 = U\big(y[G_0] - \sigma - \beta m\big)\ ,$$
$$U_1 = U\big(y[G_1] - \sigma - \beta m\big)\ .$$

The physician's utility is additive-separable in income, professional ethics and medical effort. He is risk-neutral in income because he is able to spread the risk over all patients. He receives remuneration for the supplied medical services that consist of a flat rate payment $(\omega > 0)$ and a cost reimbursement component per unit of medical treatment $(\delta + c)$, where c are the marginal costs of providing medical services and δ is a markup on these costs. The

pany that is lower than his initial income. If one considers a self-employed person, his income will be mainly determined by his productivity.

6 An additive-separable utility function states that the degree of risk-aversion of the income-dependent utility $\big(U(\cdot)\big)$ does not vary with the effort-level (a) (cf. Macho-Stadler / Pérez-Castrillo [1997], 19).

physician faces a supply-side cost sharing if only a fraction of the medical costs is reimbursed, $\delta < 0$ (cf. Ellis / McGuire [1990] and Ellis / McGuire [1993]). On the other hand, given $\delta > 0$, he gets a markup on the marginal costs of medical care. As a simplification, it is assumed that the marginal costs are equal to one.[7] Furthermore, the physician truthfully reports the medical services supplied. The provision of medical services produces disutility $C(m)$ despite the medical costs with $C' > 0$ and $C'' > 0$. We assume for the professional ethics that the expected state of the patient's health enters positively in the following expected utility function:

$$EV = [\omega + \delta m] - \varepsilon\big[pG_1 + [1-p]G_0\big] - C[m] \; . \tag{2.2}$$

The parameter $\varepsilon \in [0, 1]$ describes the intensity of the professional ethics. It corresponds to the constant marginal rate of substitution between profits and treatment benefits (cf. Ellis / McGuire [1990], 382). If the physician exhibits no ethics the parameter will take a value $\varepsilon = 0$. On the other hand if $\varepsilon = 1$ he acts as a perfect agent to the patient, i.e. that the provider values the patient's expected health benefits equivalently to his own profits. In general, a higher value of ε indicates that the expected health status has an increasing influence on the physician's decision.

The insurance company finances the health care expenditures against a premium σ. The insurance company pays for all treatment costs except the copayment βm. Insurance is supplied at actuarial fair premiums on a competitive insurance market where q is the exogenous probability of getting sick:[8]

$$\sigma = q\big(\omega + [\delta + 1]m - \beta m\big) \; . \tag{2.3}$$

Moreover, the following assumptions concerning the informational structure are made: First, both physician and patient have knowledge about the realized state of health but the insurer has not. This implies that it is not possible to write contracts contingent on the state of health. The production function, the relation between the input factors and the probability of a good state of health, is well-known ex ante but the physician and the patient cannot draw conclusions about the other one's action from the realized state of health. In detail, this means that the physician does not observe the patient's compliance and that the patient has no knowledge about the provider's actions. In contrast, the insurer can perfectly observe the amount of medical treatment and contracts are based on this observation.

7 Under the assumption of constant marginal costs for medical care equal to one it follows that the total remuneration is $\omega + (\delta + 1)m - m = \omega + \delta m$ (see equation (2.2)).

8 The possibility for the patient to influence this probability is neglected (no ex ante moral hazard). Furthermore, we only consider linear contracts for the patient and the physician.

2.2 Cooperative solution

As shown in the previous section, the health care sector is characterized by a situation of mutual asymmetric information between physician and patient. The medical services provided as well as the patient's effort (compliance) influence the value of the treatment for the two actors. This situation is known as double moral hazard (cf. e.g. Bhattacharyya / Lafontaine [1995], Cooper / Ross [1985], Demski / Sappington [1991] as well as Kim / Wang [1998]). To analyze the effects of double moral hazard we first consider the situation of complete information.

Both players choose their actions cooperatively to maximize the sum of expected utilities. The resulting first-best solution is specified by the following problem:

$$\text{(2.4)} \qquad \max_{a,m} \quad S = p(a,m)[U_1 + \varepsilon G_1] + \big(1 - p(a,m)\big)[U_0 + \varepsilon G_0] + (\omega + \delta m) - D(a) - C(m) \ .$$

Taking equation (2.3) into account the first-order conditions are:

$$\text{(2.5)} \qquad \frac{\partial S}{\partial a} : p_a\big[U_1 - U_0 + \varepsilon(G_1 - G_0)\big] = D' \ ,$$

$$\text{(2.6)} \qquad \frac{\partial S}{\partial m} : p_m\big[U_1 - U_0 + \varepsilon(G_1 - G_0)\big] + \delta = C' + \beta\big[pU_1' + (1-p)U_0'\big] + q(1 + \delta + \beta)\big[pU_1' + (1-p)U_0'\big] \ .$$

Equation (2.5) describes the patient's choice of effort. It states that the joint marginal expected utility of physician and patient on the left hand side equals the marginal disutility of a higher level of compliance on the right hand side. Equivalently, for equation (2.6) it follows that the joint marginal expected utility on the left hand side equals the marginal expected costs of medical treatment on the right hand side. The latter consists of the marginal disutility of providing medical services, the expected marginal costs of the patient's co-payment and the expected marginal costs of a higher insurance premium. The first-order conditions depend on the compliance and the medical services. The solution to this problem is denoted (a^*, m^*) and is first-best given the corresponding reaction curves.[9]

[9] In the strict sense, the solution to the first-best optimum cannot be described by the concept of 'reaction curves'. Nevertheless, we use this terminology to characterize the patient's best reaction on the physician's chosen action and vice versa.

2.3 Non-cooperative solution

In the presence of double moral hazard, physician and patient are unable to observe the other one's action. Both maximize their own expected utility without considering the interactions. Given the realized health state, the patient has no detailed information about the quality of medical services whereas the physician cannot observe the level of compliance.[10] Following the described approach the objective function of the patient's maximization problem is given by

$$\max_{a} \quad p(a,m)U_1 + \big(1 - p(a,m)\big)U_0 - D(a) \; . \tag{2.7}$$

The resulting first-order condition after rearranging the terms is:

$$p_a(a,m)(U_1 - U_0) = D' \; . \tag{2.8}$$

Equation (2.8) shows that the patient chooses his health-related effort by equating the marginal expected utility of compliance and the marginal disutility. Concerning the insurance parameters, it is of major interest how an increase in the coinsurance parameter affects the patient's compliance. By applying the implicit function rule one obtains from equation (2.8):

$$\frac{da}{d\beta} = \frac{p_a(U_1' - U_0')\left(m + \frac{\partial\sigma}{\partial\beta}\right)}{p_{aa}(U_1 - U_0) - D''} > 0 \; . \tag{2.9}$$

Given that equation (2.8) is a utility maximum for the patient the sufficient condition requires that the second-order derivative is negative. This is true for the denominator of equation (2.9). The first term is the second derivative of the probability of a good state of health multiplied with the difference in utility. Because of the concavity of $p(a, m)$, the second derivative with respect to patient's compliance and the resulting product is negative. The second term is subtracted and describes the second derivative of the patient's disutility D that is positive because disutility is a convex function of compliance. The nominator is also negative, first, because the difference of the marginal utility of income $(U_1' - U_0')$ is negative due to the concavity of the utility function. Second, the sum of medical services (m) and derivation of the premium with respect to the coinsurance parameter $(\partial\sigma/\partial\beta)$ is positive, even though the premium effect is negative.[11] The overall effect is that an increase in the coinsurance parameter leads to a higher effort level given the amount of medical

10 The resulting health status can only serve as a weak signal of the other party's action.

11 Partial differentiation of the insurance premium with respect to the coinsurance parameter results in $|\partial\sigma/\partial\beta| = qm < m$, as $q < 1$.

services. This will lead to a substitution of medical services by health-related effort because patient's compliance is relatively cheaper. Empirical studies show a price elasticity of demand for medical services of –0.2, i.e. the demand decreases as the price the patient has to pay increases (cf. Cutler / Zeckhauser [2000], 584 ff.).

The physician maximizes his expected utility with respect to medical services. The problem and the first-order condition are

$$\max_m \quad [\omega + \delta m] + \varepsilon\Big(p(a,m)G_1 + \big(1 - p(a,m)\big)G_0\Big) - C(m) \tag{2.10}$$

and

$$\delta + p_m(a,m)\,\varepsilon(G_1 - G_0) = C' \ . \tag{2.11}$$

From the physician's point of view, the optimal amount of medical services is the level where the total marginal utility consisting of the marginal utility of income and the marginal utility of a better state of health equals the marginal disutility.

In analogy to the patient one can ask how changes in the remuneration system affect the amount of medical services supplied. The physician is risk-neutral in income. Therefore, his decision does not depend on the flat rate payment but only on the reimbursement parameter (δ). This leads to the following expression:

$$\frac{dm}{d\delta} = -\frac{1}{\varepsilon\, p_{mm}(G_1 - G_0) - C''} > 0 \ . \tag{2.12}$$

The denominator is the sufficient condition for a maximum and therefore negative, so the overall effect is positive. Higher reimbursement of costs results in an incentive to expand the amount of medical services given the level of compliance.

For further analysis equations (2.8) and (2.11) are in the center of interest. They specify the actions of a player as a function of the behavior of the other player. This means that the patient chooses his effort as a reaction to the medical services and vice versa. These reaction curves specify a Nash-equilibrium, further denoted as the non-cooperative solution (a^N, m^N).

2.4 Comparison of both solutions

The effects of asymmetric information can be shown by a comparison of the full information solution and the double moral hazard situation. Under the assumption of higher marginal disutility the patient's compliance is higher in the full information case than in the case of asymmetric information. Using equations (2.5) and (2.8) one can show that $a^*(m) > a^N(m)$ if $D'(a^*) > D'(a^N)$. For this to be true it follows that the left hand side of equation (2.5) exceeds the left hand side of equation (2.8):

$$a^*(m) > a^N(m) \Leftrightarrow p_{a^*}\left[U_1 - U_0 + \varepsilon\,(G_1 - G_0)\right] > p_{a^N}[U_1 - U_0]\ . \tag{2.13}$$

This condition is fulfilled if the marginal expected utility of the compliance in the case of cooperation is higher than the marginal expected utility in the case of non-cooperation. Then the first-best level of compliance is higher than the non-cooperative level. In this case the patient realizes that the physician benefits from a higher state of health. As long as the physician reveals some professional ethics, the patient has a higher effort level in the full information case than under asymmetric information, given the level of medical services. Therefore, it follows that the first-best reaction curve runs above the one in the non-cooperative case.

For the physician one can draw the following conclusions: Suppose that the level of medical services under full information is higher than in the case of asymmetric information $\left(m^*(a) > m^N(a)\right)$. This results in a higher marginal disutility because of the convexity of the disutility function. In this case, the left hand side of equation (2.6) minus the marginal co-payment has to exceed the left hand side of equation (2.11):

$$\begin{aligned} &m^*(a) > m^N(a) \Leftrightarrow \\ &p_{m^*}\left[U_1 - U_0 + \varepsilon(G_1 - G_0)\right] + \delta - \beta\left[pU_1' + (1-p)U_0'\right] \\ &- q(1 + \delta - \beta)\left[pU_1' + (1-p)U_0'\right] \\ &> \delta + p_{m^N}\varepsilon(G_1 - G_0)\ . \end{aligned} \tag{2.14}$$

The left-hand side of the inequality condition denotes the expected marginal utilities of medical services (m^*) for physician and patient minus the expected marginal co-payment in the case of cooperation and the expected effect of a marginal increase in the insurance premium (see equation (2.6)). The right-hand side is the expected marginal utility of medical services (m^N) for a non-cooperative solution (see equation (2.11)). The amount of medical services is higher in the case of cooperation if the inequality (2.14) strictly holds, i.e. $m^* > m^N$. In particular, the strict inequality always holds if the exogenous probability of getting sick q is not too large. In this situation, the reaction curve of the physician runs above the non-cooperative one.

3. The relation between patient's compliance and medical services

3.1 Substitutes vs. complements

The outcome of the health production process is influenced by the choice of medical services and the patient's effort that both affect the probability of a good state of health. The input levels are determined by the reaction curves. For the following analysis it is assumed that the reaction curves form a unique and stable Nash-equilibrium. Up to now, only the difference between the reaction curves of full and asymmetric information has been considered. At this point, it is necessary to analyze the effects of the reaction curves' slopes on the levels of medical services and compliance. Especially, the connection of the two input variables and their interdependences deserve further discussion. Therefore, the relation of the actions of both players under asymmetric information lies in the center of interest.

In the subsequent analysis it is assumed that the action of one player is strictly positive $(a, m > 0)$ even if the other party is providing no input at all. Otherwise, no equilibrium in positive actions exists (cf. Cooper/Ross [1985], 106). In termini of the model this means that $a(m = 0) > 0$ and $m(a = 0) > 0$. To derive the slope of the reaction curves it is necessary to transform the first-order conditions of both the patient and the physician (equations (2.8) and (2.11)) by applying the implicit function rule. For the patient it follows that the relation between compliance and medical services is:

$$\frac{da}{dm} = \frac{\beta p_a(U_1' - U_0') - p_{am}(U_1 - U_0)}{p_{aa}(U_1 - U_0) - D''} \quad . \tag{3.1}$$

The denominator is the sufficient condition of the patient's utility maximum and therefore negative. The sign of the nominator depends on two factors: The first term is only different from zero if the coinsurance rate (β) is positive (*coinsurance component*). The sign of the *strategic component* (second term) depends on the sign of the cross derivation of the probability of a good state of health (p_{am}). This derivation shows how the marginal productivity of the compliance changes due to an increase in medical services.[12] The expressions *strategic complements* and *strategic substitutes* will be used for this behavior (cf. Bulow et al. [1985], 494). In difference to the common concept of complements and substitutes, which describes a direct relation between variables, the strategic concept illustrates the effect of one variable on the marginal product of the other one. This means that e.g. additional medical services raise the

12 It is worth mentioning that an increase in one of the input factors increases the probability of a good state of health so that the term 'marginal productivity' is not correct in a strict sense.

marginal productivity of the health-related effort. Thus, the probability of a good state of health increases (strategic complements). In the reverse case, a higher level of compliance lowers the marginal productivity of medical services (strategic substitutes). An example for the former effect is obeying a therapeutic advice whereas strategic substitutes are present if the patient does not visit his physical therapist but practices the exercises on his own.

From equation (2.11) one obtains for the slope of the physician's reaction curve:

$$\frac{dm}{da} = -\frac{\varepsilon p_{am}(G_1 - G_0)}{\varepsilon p_{mm}(G_1 - G_0) - C''} \ . \tag{3.2}$$

Here the denominator also corresponds with the sufficient condition for the physician's utility maximum and is therefore negative. A closer look at the nominator shows that the slope of the reaction curve only depends on the existence of strategic complements or strategic substitutes. The absence of both strategies is called *strategic independence*. For a physician with no professional ethics ($\varepsilon = 0$) the slope is always zero.

3.2 Results without coinsurance

Without considering the effects of a co-payment for the patient the slopes of the reaction curves only depend on the three strategic effects, i.e. whether the inputs are strategic substitutes, complements or independent. Equation (3.1) for the patient then reduces to

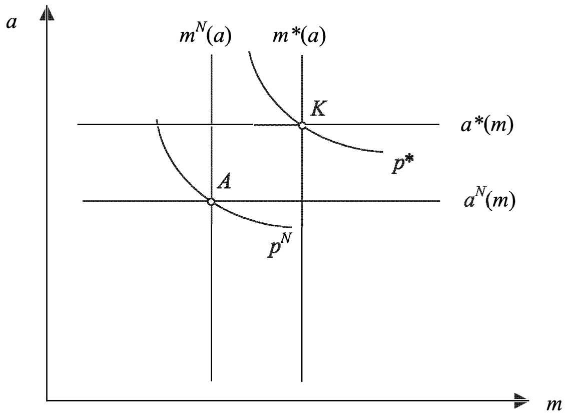
$$\frac{da}{dm} = -\frac{p_{am}(U_1 - U_0)}{p_{aa}(U_1 - U_0) - D''} \ . \tag{3.3}$$

whereas equation (3.2) for the physician remains unchanged:

$$\frac{dm}{da} = -\frac{\varepsilon p_{am}(G_1 - G_0)}{\varepsilon p_{mm}(G_1 - G_0) - C''} \ . \tag{3.4}$$

The denominator of both equations is still negative. If the marginal productivity is independent of the level of the other input ($p_{am} = 0$) both reaction curves are inelastic to changes in the other input ($da/dm = 0$ and $dm/da = 0$). Figure 1 illustrates this case.

The axes show the level of medical services (m) and the patient's effort level (a). The reaction curves a^* and m^* indicate the first-best solution under full information. Compared to the solution in the case of asymmetric information (a^N and m^N) one recognizes that for a given amount of a or m respectively, the chosen action of the other input is on a higher level. Point K indicates the first-best solution that lies in the point of intersection of the corresponding reaction curves. Compared to the non-cooperative equilibrium (point A) the level of medical services as well as the compliance is higher for cooperation. The iso-probability curve p^* denotes the probability of a good state of health under full information, p^N signifies this probability in the case of asymmetric information. The negative slope results from total differentiating $p(a, m)$.[13] Curves further off the origin show higher probabilities for a good state of health and therefore, at least the level of one input increases (cf. Lanoie [1991]).



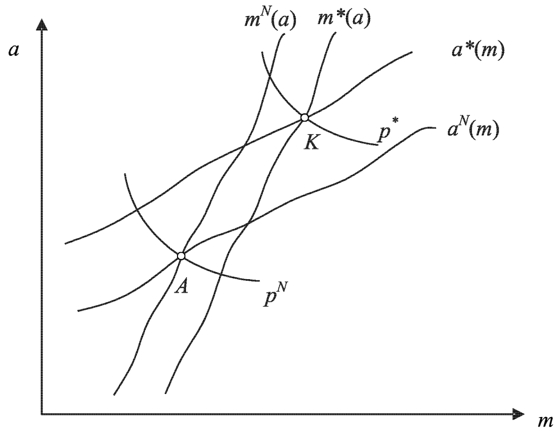
Source: Following Cooper / Ross, 1985.

Figure 1: Independence of compliance and medical services

If an increase in the medical services raises the marginal productivity of the patient's compliance or if higher health-related effort leads to a higher marginal productivity of medical services we face the situation of *strategic complements* ($p_{am} > 0$). Here, the nominator of equations (3.3) and (3.4) is positive and so are the slopes of the reaction curves: $da/dm > 0$ and $dm/da > 0$ (see figure 2).[14] The reaction curves are no straight lines anymore. The actual slopes depend on the utility of income, the disutility of compliance and on the probability of a recovery.[15]

13 Given the function $p(a, m)$, it is necessary for the convexity of the iso-probability curves that the Hessian matrix is negative semidefinite. This applies if $p_{aa} p_{mm} - p_{mm}^2 > 0$. In the case of strategic independence, the iso-probability curve is always convex. For strategic complements and substitutes the squared cross derivative of $p(\cdot)$ has to be smaller than the product of the second order partial derivatives.

14 As a simplification, in the following presentation the reaction curves do not need to be linear in the other input. Actually, this depends on the utility function as well as on the density function of the probability of a better health status.



Source: Following Cooper / Ross, 1985.

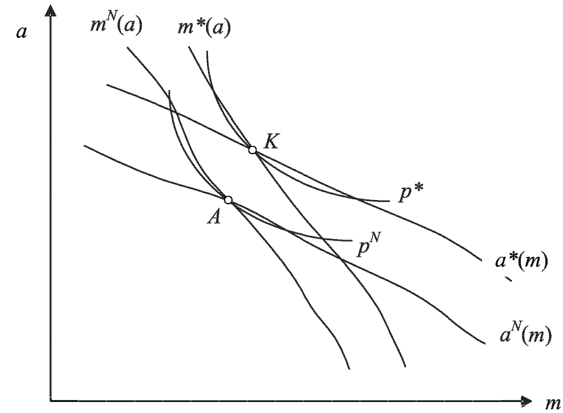
Figure 2: Compliance and medical services in the case of strategic complements

The lines a^* and m^* represent the full information reaction curves. The resulting first-best equilibrium is again denoted by point *K*. In comparison to the non-cooperative solution, given in point *A*, it is obvious that in the situation of double moral hazard both the amount of medical services and the patient's compliance are at a lower level. This result can be regarded as an incentive problem because of the lack of knowledge about the consequences of individual actions.

In the case of *strategic substitutes,* the marginal productivity of the compliance falls due to an increase in the medical services and vice versa. An example for the latter case is a slower recovery through exaggerated exercises. The slope of the reaction curves is negative ($da/dm < 0$ and $dm/da < 0$) because the nominator of equations (3.3) and (3.4) is negative. While the results in the case of strategic independence and complements are unambiguous, a more sophisticated analysis is necessary when strategic substitutes exist (see figures 3 and 4).

In the first case (figure 3), the non-cooperative reaction curves are below the ones in the full information case. The asymmetric information results in lower levels of both inputs. The iso-probability curve p^N is beneath p^* which means that the probability to recover from an illness is lower if double moral hazard is present. These results correspond to the results in the case of strategic complements.

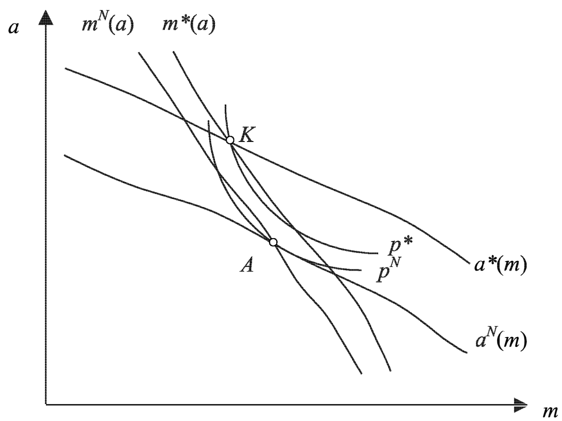
[15] In the following figures, the reaction curves are drawn as increasing respectively decreasing lines with a varying slope.



Source: Following Cooper / Ross, 1985.

Figure 3: Comparable effects in the case of strategic substitutes

A different result is shown in figure 4. In this case, the reaction curves under asymmetric information again run below the first-best reaction curves. However, on the one hand the level of compliance is lower where on the other hand one obtains an increase in medical services (cf. points *K* and *A*). The explanation is quite intuitive (cf. Cooper / Ross [1985], 108 f. and Yavaş [1995], 253 f.): Starting from the cooperative solution one first observes decreasing levels of both inputs as an effect of the mutual asymmetric information.



Source: Following Cooper / Ross, 1985.

Figure 4: Different effects in the case of strategic substitutes

This corresponds to the shift in the reaction curves. A second effect occurs because the reduction of the patient's compliance results in a higher marginal productivity of medical services for a good state of health. Therefore, the physician will increase his treatment services. The same is true for the reversed case in which a reduction of the medical services leads to a higher marginal productivity of the compliance. It is possible that the second effect of a higher marginal productivity outweighs the first effect of asymmetric information for either the patient or the physician with the consequence that the first-best level is exceeded.

3.3 Introduction of a coinsurance

The introduction of a coinsurance for the patient affects the resulting Nash-equilibrium. First, the coinsurance component in equation (3.1) now differs from zero. Moreover, the utility and the marginal utility of the patient are influenced by his medical expenditures conditional on the coinsurance parameter. Equation (3.1) indicates that the slope of the patient's reaction curve depends on the characteristics of the *coinsurance component* as well as the *strategic component.*

In addition, the introduction of a coinsurance leads to a variation in net income, whereby the utility as well as the marginal utility are influenced.[16] Generally, using equation (2.9) one obtains the effect that c. p. a higher coinsurance rate leads to an increase in the health-related effort. The exact influence of a coinsurance depends on the difference in patient's utility and on the difference in the marginal utilities in the nominator (cf. equation (3.1)). First, the difference in utility resulting from a good health status and utility from a bad health status $(U_1 - U_0)$ is always positive. Because of the concavity of the utility function, this difference increases if a coinsurance is implemented. If one neglects the influence of the coinsurance component, this would result in an increase of both the nominator and the denominator where the relative increase is higher for the nominator. Second, the difference in marginal utilities $(U_1' - U_0')$ is negative. The overall effect depends on the influence of the coinsurance component and the sign of the strategic component.[17]

The introduction of a coinsurance leads to a positive slope of the patient's reaction curve in the case of *independence* of compliance and medical services $(p_{am} = 0)$ (dotted line a^{β} in figure 5). Here, the strategic component equals zero and the coinsurance effect is negative. Hence, from the patient's point of view a complementary relation between the physician's medical services and his own compliance is present.

16 An introduction of a coinsurance rate leads to a lower insurance premium so that the patient's net income is not reduced by the full amount of the co-payment.

17 The derivation of the comparative statics is presented in the appendix.

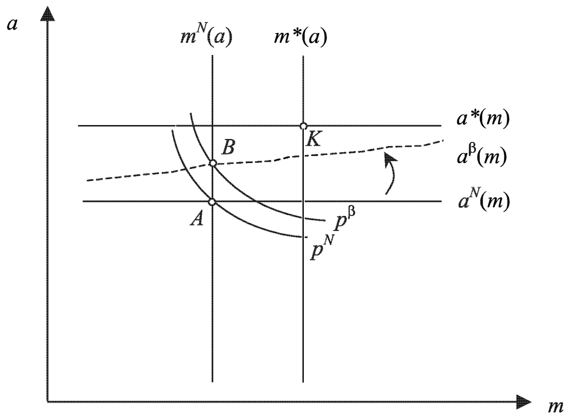


Figure 5: Independence and coinsurance

In contrast to the situation without coinsurance, an increase in the amount of medical services raises the compliance. This leads to a new equilibrium (point B) on a higher iso-probability curve (p^β). For each level of the physician's services, the patient will raise his health-related effort because it is now relatively cheaper in comparison to the medical services. In addition, this higher level leads to an increase in the patient's expected utility of income.

Given *strategic complements* ($p_{am} > 0$), the coinsurance component remains negative whereas the sign of the strategic component is positive. The difference in both effects is negative and the nominator increases more than the denominator compared to the situation without coinsurance. Overall, the slope of the patient's reaction curve is steeper and a coinsurance leads to an amplification of the results without coinsurance. The patient's reaction curve runs above the one in the case of non-cooperation. It follows that the patient will further increase his compliance compared to the situation without a coinsurance for a given level of medical services. The dotted line (a^β) in figure 6 presents this effect.[18]

Line a^β specifies the new reaction curve for the patient whereas the physician's one remains unchanged. It is obvious that in the intersection with the reaction curve m^N (point B) the level of medical services as well as the level of compliance are higher than in the case without coinsurance. Thereby the probability of a recovery (p^β) increases compared to the one in the Nash-equilibrium without coinsurance (p^N).

---

18 The real slope of the reaction curve a^β is irrelevant for the qualitative result. The presented change of the reaction curve in the following figures exemplifies the effect of the coinsurance.

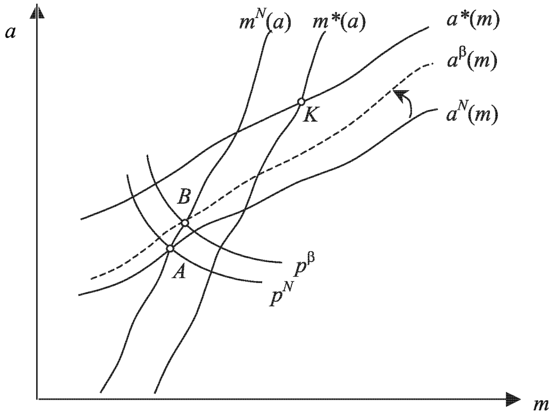


Figure 6: Strategic complements and coinsurance

One obtains different results for the case of *strategic substitutes* $(p_{am} < 0)$. Here, the implementation of a coinsurance possibly reduces the slope (in absolute terms) of the patient's reaction curve (see figure 7). In contrast to the situations with strategic independence and complements, here, the coinsurance and the strategic component have different effects. For the strategic component, the increasing difference in utility leads to a steeper reaction curve. The effect of the coinsurance component is a flatter reaction function. If the strategic effect dominates the coinsurance effect, i.e. if $-p_{am}(U_1 - U_0) + \beta p_a(U_1' - U_0') > 0$, the nominator is still positive but decreases. Given the physician's treatment decision the patient's compliance is c. p. at a higher level. In principle, it is possible that the influence of the coinsurance component prevails against the strategic component so that the slope of the reaction curve is positive.

It is obvious that because of the new reaction curve (a^β) the amount of medical services decreases whereas the amount of compliance increases (point B). It is also possible that the patient's effort is above the first-best level. Thus, compliance is substituted for medical services.

The slope of the iso-probability curve given by the marginal rate of substitution between compliance and medical services depends on the marginal productivity of both inputs. For the case of a steeper iso-probability curve, it is possible that the probability of a good state of health decreases. This is shown as the movement from point A to point B in figure 8. The higher level of compliance and the associated increase in the probability cannot compensate the decreasing probability of a good state of health due to the decreasing medical services.

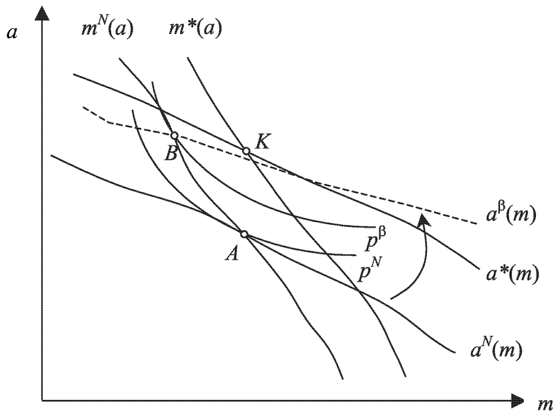


Figure 7: Strategic substitutes and increasing probability of recovery

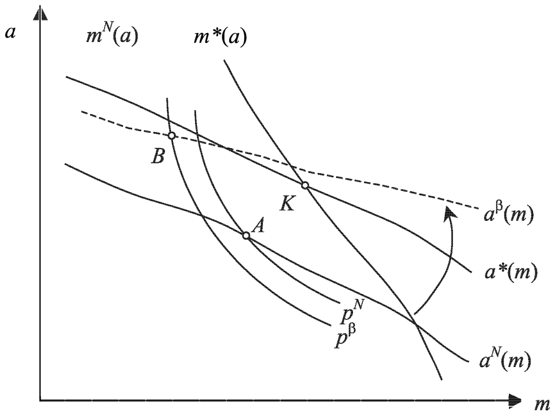


Figure 8: Strategic substitutes and decreasing probability of recovery

In summary, one can state that for all forms of strategic interaction the compliance increases due to an introduction of a coinsurance. The explanation is that the health-related effort and the medical services are productive inputs. Without coinsurance, the marginal costs of the consumption of medical services are zero. After the implementation, the patient faces positive treatment costs. Compliance becomes relatively cheaper in comparison to medical services.

4. Welfare and policy implications

The introduction of a coinsurance gives interesting insights into the physician-patient-relationship and the individual behavior. From an allocative perspective, different questions about the resulting welfare aspects arise. If an implementation of a coinsurance leads to a welfare improvement for the patient, the following condition has to be fulfilled:

$$(4.1) \qquad p^{\beta} U_1^{\beta} + (1 - p^{\beta}) U_0^{\beta} - D^{\beta} > p^N U_1^N + (1 - p^N) U_0^N - D^N \ .$$

The superscript N denotes the Nash-equilibrium without coinsurance, the superscript β the situation after the implementation of a coinsurance. The left-hand side of the inequality is the patient's expected utility in the case of a coinsurance. If this expression exceeds the right-hand side of the inequality (the patient's expected utility without coinsurance) out-of-pocket payments lead to a welfare improvement.

To simplify the analysis the patient's utility is divided into three pieces: the probability of a recovery (p), the expected utility of income in a good and a bad health status (U_1 and U_0) and the disutility of compliance (D). Beginning with the last element it is clear from equation (2.9) that an increase in the coinsurance rate increases the compliance. Moreover, because disutility is a convex function of compliance the introduction of a coinsurance results in a higher disutility. Next, the impact of a coinsurance on net-income is important. It is obvious that there is a direct reduction of income. The indirect effect results in a lower insurance premium. The net effect is that the income after a medical treatment is lowered by the coinsurance but not at the full amount. Overall, the income effect leads to a decrease in the utility of a good and a bad health status. The last effect is the probability of a recovery. The consequence of patient's compliance is a higher probability but the entire outcome depends on the medical services that are influenced by the level of compliance.

The total effect remains unclear because out-of-pocket payments reduce the net income and increase the disutility resulting from a higher level of compliance whereas the impact on the probability of a recovery depends on the strategic interaction. A welfare improvement depends on the magnitude of the probability of a recovery which itself depends on the strategic interaction between compliance and medical services. Starting with strategic independence of health inputs the introduction of a coinsurance results in an increase in the level of compliance with no change in the level of medical services. This increases the probability of a recovery and the possibility of a welfare improvement. If one takes a look at the effects of strategic complements both the levels of compliance and medical services rise. Therefore, the increase in the probability p as well as the likelihood of a welfare improvement is higher compared to independence. The situation is different given strategic substitutes. Here, we face the problem that an increase as well as a decrease of the probability of a recovery is possible. In the first case, the reduction in medical services is compensated by an increase in the level of compliance. Nevertheless, the increase in the probability is lower than in the other cases of strategic interaction and welfare improvement becomes less likely. In the second case, a rise in compliance cannot compensate for the reduction of medical services and the probability of a recovery decreases making a welfare improvement almost impossible.

To summarize the findings it is worth mentioning that the introduction of compliance might improve patient welfare. This result crucially depends on the form of strategic interaction between the patient's compliance and the physician's medical services. A global coinsurance for all kinds of illness neglects these findings and leads to disincentives with respect to health production. Therefore, health politics cannot solve the problem of rising health expenditures with the single instrument of a global coinsurance. Instead, there is a need for illness-specific out-of-pocket payments that on the one hand stimulate patient's compliance and on the other hand keep incentives for providing the necessary medical care. Beyond this, two points exist that one has to keep in mind. First, the analysis abstracts from the decision about the physician contact. Second, out-of-pocket payments have an influence on this decision that is not incorporated in the presented model. A coinsurance is welfare reducing if the patient decides not to visit a doctor and if this choice lowers the probability of a recovery.

5. Conclusion

The paper at hand analyzes how the relationship between the medical services provided by the physician and the patient's compliance influences the resulting equilibrium at the point of treatment. The utilization of medical services and the level of health-related effort strongly depend on their productivity characteristics, i.e. if the inputs are strategic complements, substitutes or independent factors. In contrast to the other solutions, the non-cooperative equilibrium in the case of strategic substitutes does not necessarily lead to lower medical services and compliance compared to the first-best solution. It is possible that either the level of medical services or compliance is above the first-best level while the other input level is below. The introduction of a coinsurance for the patient alters these findings. The patient's reaction curve is shifted towards the cooperative one. This implies that the patient always chooses a higher level of compliance for a given level of medical services. While in the case of independence the amount of medical services remains unchanged, in the case of strategic substitutes the latter is reduced. If we consider strategic complements both the effort and the supply of medical services increase.

The findings are interesting with regard to a possible welfare improvement. Two conditions must be fulfilled: First, as a necessary condition the difference between the utility of income after introducing a coinsurance in a good state of health and the expected utility of income without coinsurance must exceed the difference in disutility. Second, the probability of a recovery has to be large enough that the expected utility in the case of a coinsurance is larger than the expected utility without. In particular, the welfare effect depends crucially on the strategic interaction of compliance and medical services. In the case of strategic complements a welfare improvement is more likely than in the case of independence. For strategic substitutes it is possible that the gain in patient's compliance is not sufficient for a compensation of the reduction of medical services and the probability of a recovery declines.

Like in other models of asymmetric information the assumptions made concerning the knowledge of physician and patient are problematic in different ways. It is not clear whether physician and patient really know about the strategic interactions and whether they realize the consequences of their individual actions. If the patient has no information about the effects of a higher level of compliance on the health production process, the ex ante decisions of the patient might prove wrong ex post. Therefore, it can be considered that the physician communicates the necessary information about the strategic interaction of medical services and compliance to the patient.

Appendix: Comparative statics of an introduction of a coinsurance

The derivative of equation (3.1) with respect to the coinsurance parameter β is taken to obtain the effect of the introduction of a coinsurance. Therefore, we apply the quotient rule:

(A.1) $$\frac{d(da/dm)}{d\beta} = \frac{f'(\beta)g(\beta) - f(\beta)g'(\beta}{g^2(\beta)} ,$$

where $f(\beta)$ is the nominator and $g(\beta)$ is the denominator of equation (3.1) and f' and g' are the partial derivatives with respect to β:

(A.2) $$f'(\beta) = p_a(U_1' - U_0') + \beta m p_a(U_0'' - U_1'') + m p_{am}(U_1' - U_0') ,$$

(A.3) $$g(\beta) = p_{aa}(U_1 - U_0) - D'' < 0 ,$$

(A.4) $$f(\beta) = \beta p_a(U_1' - U_0') - p_{am}(U_1 - U_0) ,$$

(A.5) $$g'(\beta) = p_{aa} m(U_1' - U_0') > 0 .$$

The overall effect of a coinsurance depends on the sign of equation (A.1). The denominator is positive, so the effect depends on the sign of the nominator. The effect of a higher coinsurance is unambiguous for strategic complements and strategic independence. Here, (A.2) and (A.4) are negative and the derivative of da/dm with respect to β is positive which means that the introduction of a coinsurance results in a steeper reaction curve for the patient. Therefore, for all levels of medical services, the patient exhibits a higher level of compliance.

If a and m are strategic substitutes, the effect of a coinsurance is complex. As before, if (A.2) and (A.4) are negative, we find that the patient's reaction curve is flatter, implying that a higher level of medical services results in a higher level of compliance. The effect changes if (A.2) and (A.4) are positive. Then the negative slope of the patient's reaction curve increases and it becomes steeper. For (A.4) to be positive, the coinsurance component has to exceed the strategic component. The term (A.2) is only positive if the third term exceeds the first two. If one of them, (A.2) or (A.4), is negative, then it is possible that the overall effect of a higher coinsurance rate is positive or negative. In order to simplify the analysis in section 3.3 we further assume that in the case of strategic substitutes a higher coinsurance rate leads to a flatter reaction curve.

References

Arrow, K. J. (1985): The Economics of Agency, in: J. W. Pratt / R. J. Zeckhauser (eds.), Principals and Agents: The Structure of Business, Boston, 37 – 51.

Bhattacharyya, S. / *Lafontaine,* F. (1995): Double-Sided Moral Hazard and the Nature of Share Contracts, Rand Journal of Economics, 26, 761 – 781.

Blomqvist, A. (1991): The Doctor as Double Agent: Information Asymmetry, Health Insurance, and Medical Care, Journal of Health Economics, 10, 411 – 432.

Bulow, J. I. / *Geanakoplos,* J. D. / *Klemperer,* P. D. (1985): Multimarket Oligopoly: Strategic Substitutes and Complements, Journal of Political Economy, 93, 488 – 511.

Börsch-Supan, A. H. (1998): Anreizprobleme in der Renten- und Krankenversicherung, in: R. Ackermann / S. Cassel / E. Denner (eds.), Offen für Reformen?: Institutionelle Voraussetzungen für gesellschaftlichen Wandel im modernen Wohlfahrtsstaat, Baden-Baden, 271 – 290.

Cooper, R. / *Ross,* T. W. (1985): Product Warranties and Double Moral Hazard, Rand Journal of Economics, 16, 103 – 113.

Cutler, D. M. / *Zeckhauser,* R. J. (2000): The Anatomy of Health Insurance, in: A. J. Culyer / J. P. Newhouse (eds.), Handbook of Health Economics Vol. 1A, Amsterdam et al., 563 – 643.

Demski, J. S. / *Sappington,* D. E. M. (1991): Resolving Double Moral Hazard Problems with Buyout Agreements, Rand Journal of Economics, 22, 232 – 240.

Dionne, G. / *Contandriopoulos,* A. P. (1985): Doctors and Their Workshops: A Review Article, Journal of Health Economics, 4, 21 – 33.

Ellis, R. P. / *McGuire,* T. G. (1990): Optimal Payment Systems for Health Services, Journal of Health Economics, 9, 375 – 396.

– (1993): Supply-Side and Demand-Side Cost Sharing in Health Care, Journal of Economic Perspectives, 7, 135 – 151.

Gaynor, M. (1994): Issues in the Industrial Organization of the Market for Physician Services, Journal of Economics and Management Strategy, 3, 211 – 255.

Holmström, B. (1979): Moral Hazard and Observability, Bell Journal of Economics, 10, 74 – 91.

Kim, S. K. / *Wang,* S. (1998): Linear Contracts and the Double Moral-Hazard, Journal of Economic Theory, 82, 342 – 378.

Kortendieck, G. (1993): Gesundheitsökonomie und Wirtschaftspolitik: Neoklassische versus österreichische Markttheorie dargestellt am Beispiel des Gesundheits- und Krankenversicherungswesens, Freiburg i. Br.

Lanoie, P. (1991): Occupational Safety and Health: A Problem of Double or Single Moral Hazard, Journal of Risk and Insurance, 58, 80 – 100.

Ma, C. A. / *McGuire,* T. G. (1997): Optimal Health Insurance and Provider Payment, American Economic Review, 87, 685 – 704.

Macho-Stadler, I. / *Pérez-Castrillo,* J. D. (1997): An Introduction to the Economics of Information: Incentives and Contracts, Oxford University Press, Oxford and New York.

Pauly, M. V. (1980): Doctors and Their Workshops: Economic Models of Physician Behavior, Chicago, London.

Schneider, U. (2002): Theorie und Empirie der Arzt-Patient-Beziehung: zur Anwendung der Principal-Agent-Theorie auf die Gesundheitsnachfrage, Frankfurt et al.

Selden, T. M. (1990): A Model of Capitation, Journal of Health Economics, 9, 397 – 409.

Shmanske, S. (1996): Information Asymmetries in Health Services: The Market Can Cope, Independent Review, 1, 191 – 200.

Wille, E. / *Ulrich,* V. (1991): Bestimmungsfaktoren der Ausgabenentwicklung in der gesetzlichen Krankenversicherung (GKV), Hansmeyer, K. (ed.), Finanzierungsprobleme der sozialen Sicherung II, Berlin, 9 – 104.

Yavaş, A. (1995): Seller-Broker Relationship as a Double Moral Hazard Problem, Journal of Housing Economics, 4, 244 – 263.

Zweifel, P. (1994): Eine Charakterisierung von Gesundheitssystemen: Welche sind von Vorteil bei welchen Herausforderungen? in: Oberender, P. (ed.), Probleme der Transformation im Gesundheitswesen, Baden-Baden, 9 – 43.

Zweifel, P. / *Lehmann,* H. / *Steinmann,* (2001): Patching up the Physician-Patient Relationship: Insurers vs. Governments as Complementary Agents, Paper presented at the Annual Meeting of the German Economic Association (Verein für Socialpolitik), 25. – 28. September, Magdeburg.