Financing Medical Specialist Services in The Netherlands: Welfare Implications of Imperfect Agency

Kees Folmer and Ed Westerhout

Abstract

From 1995 onward the financing scheme for medical specialist services in the Netherlands has moved from a fee-for-service scheme to a capitation scheme. This paper analyzes the economic and welfare effects of this policy change. The paper adopts a numerical model that integrates demand and supply considerations and recognizes the potential roles of moral hazard and supplier-induced demand. The paper finds that the shift in financing regime has been welfare-reducing. The policy change induced medical specialists to lower the supply of health services which was already too low from a welfare point of view. This finding is robust to significant changes in major parameter values.

Keywords: fee-for-service scheme, lump-sum budget, medical specialists, moral hazard
JEL-codes: D60, H21,I18

INTRODUCTION

For many years, the services delivered by medical specialists in the Netherlands were financed according to a fee-for-service scheme. Although the scheme in place was frequently adjusted, it basically linked the income of medical specialists to the volume of their output. This financing scheme was heavily debated. In particular, it was argued that it induced medical specialists to deliver more services than was in the patients' interests, that it increased aggregate expenditure on specialist care and made costs uncontrollable at the macro level.

In 1995 things have begun to change. The Dutch government allowed yearly negotiations between health insurers, hospitals and the medical specialists affiliated with these hospitals, to negotiate about a budget for medical specialists independent of the volume of their services.

This paper investigates the economic and welfare effects of the introduction of this capitation system. We adopt a principal-agent approach in which medical specialists take consumption decisions. Demand enters these decisions as specialists incur a utility loss if they deviate from the preferences of their patients. By including both demand and supply, we do justice to empirical evidence that shows that medical consumption is responsive not only to demand factors like the out-of-pocket price of medical consumption and patient income (e.g. [1-2]) but also to supply factors like physician income and the number of physicians relative to the population (e.g. [3-5]). Our model is part of a larger one that covers almost the complete Dutch health care sector (see [6] for a description of the behaviour of physicians and hospitals and [7] for a description of pharmaceutical markets).

Our analysis indicates that the financial reform has led to a decline in the consumption of medical specialist services. Furthermore, it finds that patients were hurt by this decline in medial consumption, whereas medical specialists benefitted from the concomitant increase in leisure time. On net, society suffered a welfare loss. The reason is that due to (negative) supplier-induced demand, medical consumption was already too low before the financial reform. The reform aggravated this inefficiency by inducing medical specialists to reduce supply, thereby lowering medical consumption still further. These results are robust. Indeed, the sensitivity analysis undertaken indicates that significant changes in the parameter configuration of our model do not affect our qualitative results.

The numerical configuration of the model is important however as theory alone cannot sign the welfare effect of the reform. One reason is that the relation between initial demand and supply is unclear. Therefore, a reduction of supply brought about by the reform may increase or decrease the gap between demand and supply. A second reason is that moral hazard considerations lend support to a fee reduction of a certain size. The fee reduction that is implied by the reform may as well be too large or too small. Thus, neither the supplier-induced demand distortion or the moral hazard distortion present in our model can be used to conclude beforehand what will be the sign of the welfare effect. Our conclusion that the welfare effect was negative thus depends on the numerical configuration of the model.

Our approach fits into the literature on the welfare effects of medical insurance (e.g. [8-10]), but extends it by recognizing the independent role of physicians. This extension has an impact on the relation between fee and volume changes as well as on the welfare consequences of the former, as fee changes generally also affect

the well-being of physicians. Next, our paper joins the literature on physician responses to fee changes (e.g. [11-12]), but extends it by including the role of patients. This allows us to demonstrate that supplier-induced demand may help to combat the moral hazard that is due to insurance (see also [13]). Closest to our paper are Ellis and McGuire [14-15] which also focus on the interaction between demand and supply considerations. Different from these two papers is that our paper explores the effects of a real-world experiment and adopts specifications for the demand and the supply of health care that have a strong empirical base. Furthermore, our paper does not take demand and supply policies as independent, but as related, due to the constraint that the financial reform that is analyzed should not change the income of physicians in the aggregate.

The setup of our paper is as follows. First, it describes some features of the reform of the financing scheme. Next, it constructs our model which is subsequently used to calculate the effects of the financial reform. Finally, it provides some concluding remarks.

THE FINANCIAL REFORM

In 1995, medical specialists, hospitals and health insurers in a number of regions started to experiment with schemes in which medical specialists were given a lumpsum budget. The specialists were free to participate in these so called local initiatives. However, if they chose not to participate, they would run the risk of ongoing fee reductions as part of a policy of macro budgeting. Tired of political conflicts and faced with the prospect of falling incomes, most medical specialists opted for participation in the new financing scheme.

Although the new schemes have removed the direct link between actual medical production and the income of the individual specialist, the reform did not produce substantial income effects, as on average budgets were set such as to keep specialist income on its pre-reform level. This does not imply that individual physicians did not face important shifts in their incomes. Indeed, reduction of income differentials across different specialties was also an important element that was negotiated in the local initiatives.

An analysis of the new financing scheme in five hospitals involved in the first round of experiments indicates that it has changed the behavior of medical specialists. The new system has affected output variables like the probability of an admission and the average waiting time ([16]). In addition, a statistically significant effect has been found for the referral ratio: after the introduction of the new financing scheme, a significantly larger fraction of patients of medical specialists was referred back to a general practitioner ([17-18]). On the other hand, no significant effects upon the average duration of stay were found. This could reflect that the duration of stay is to a large extent driven by technological growth.

Aggregate time series data on hospital admissions, outpatient treatments and outpatient visits point in the same direction ([19]). Hospital admissions grew 1.1 percent a year during the period 1990-1994, but with the onset of the new financing regime the average yearly rate of growth dropped to -1.4 percent (period 1995-2000). Outpatient treatments grew firmly during 1990 -1994 at an average yearly rate of 10%, but the yearly growth rate has declined since 1995 to 5.5%. The development of outpatient consultations yields the same evidence. An average rate of increase during 1990-1994 of 1.3% was followed by an average decline in later years: -0.1% during 1995-2000.

In the period from 1995 to 2000, the insured population grew at rates of 0.5 percent per year, whereas the share of people aged 65 and older increased from 13.2 to 13.5 in this period. Apart from the reform under discussion, no other major policy reforms took place in the same period . This suggests that the developments in hospital production during the years 1995-2000 are to be attributed to the change in the financing scheme for medical specialists, at least for some part.

Summing up, micro evidence on specialist behavior in the hospitals that were the first to experiment with the new financing scheme points to a negative effect on the volume of specialist services. Macroeconomic time series on various output variables in the 1995-2000 period hint in the same direction. Further, it is useful to remark that currently a new financing scheme, in particular a scheme based upon diagnosis-related-groups, is under discussion in the Netherlands, mainly because it is felt that the current lumpsum budgeting system offers too little incentives for specialist production.

A MODEL OF MEDICAL CARE

Medical specialists

Basically, the model that describes the behavior of medical specialists is a neoclassical labor supply model. However, we extend it with ethical costs in order to impose an agency relationship between the patient and the physician.

In principle, the presence of imperfect information at the side of the patient makes it possible for the medical professional to deviate from the patient's interest and pursue other (personal) objectives. On the other hand, we do not think it reasonable to assume that the physician can fully neglect the patients' interests. Reputation considerations, the medical oath, or medical ethics all lead us to regard physicians as agents of patients. This special agency-feature is allowed for in the model by imposing a (fixed) ethical cost on the professional when she deviates from the patient's best interest, *i.e.* the action that the patient would have selected under perfect information.

A physician k derives utility, v_k, from leisure, l_k, and the consumption of other goods and services, d_k. The utility function is specified in CES-format,

$$v_k(d_k, l_k, e_k) = \left(\alpha d_k^{-\rho} + l_k^{-\rho}\right)^{-1/\rho} - e_k \qquad \rho > -1 \tag{1}$$

where e refers to the ethical cost which is imposed on the physician when she fails to act in the patient's best interest, i.e when she fails to deliver demand. The fixed nature of the ethical costs implies that the physician has to choose between two options, viz. an ethical with zero ethical costs and a financial option which corresponds to fixed positive ethical costs. For a discussion on the fixed nature of ethical costs in our model, see [6].

The physician maximizes her utility subject to a time constraint and a budget constraint. The time constraint says that leisure time is determined from the difference between the total time allotment T and the time allocated to producing services. Denoting the supply of services by s, we have,

$$l_k = T - \mu s_k \tag{2}$$

where μ measures units of time needed to produce one treatment.

Consumption equals the physician's income, which is calculated as the revenue from services and an additional income component beyond the control of the physician. Or,

$$d_k = ts_k + h \tag{3}$$

where t stands for the fee per service and h is the physician's lumpsum income.

To derive the physician's behavior, it is useful to first obtain the solution that maximizes the physician's utility from consumption and leisure (*i.e.* the first term at the right hand side of (1)). This solution characterizes the financial option. The interior solution of the maximization problem, denoted $\bar{s}^*$, reads as

$$\bar{s}^* = \beta\frac{T}{\mu} - (1-\beta)\frac{h}{t} \tag{4}$$

where the auxiliary variable β is defined as

$$\beta = \frac{t^{\sigma-1}\mu^{1-\sigma}\alpha^{\sigma}}{1 + t^{\sigma-1}\mu^{1-\sigma}\alpha^{\sigma}} \qquad (5)$$

and $\sigma=1/(1+\rho)$ is defined as the elasticity of substitution between consumption and leisure. Equation (5) shows that $0 < \beta < 1$.

As labor supply cannot be negative, under the financial option it follows the interior solution to the corresponding maximization problem, $\bar{s}^*$, except when $\bar{s}^* < 0$, in which case 0 replaces the interior solution. Formally, $\bar{s} = \max(0, \bar{s}^*)$. The solutions for nonmedical consumption and leisure that correspond to the financial option read as $\bar{d}$ and $\bar{l}$ respectively, and follow from substituting the value of $\bar{s}$ into expressions (3) and (2).

We stress that expression (4) implies that financial physicians may supply more or less medical services than demanded, depending on the physician's particular preferences for consumption and leisure and her incentives to supply medical services. In particular, the physician may supply more than demand to raise her income but less than demand to gain leisure time.

Under the ethical option, a specialist accommodates her supply of services, $\tilde{s}$, to the corresponding demand, Z/N_s, where Z denotes aggregate demand and N_s denotes the number of medical specialists. The only exception occurs when this behaviour would imply negative leisure. In that case, T/μ replaces Z/N_s. This effect does not occur in any of our simulations, however. Apart from this exceptional case, patients receive the amount of medical services they would have demanded in the absence of any information imperfections. We use $\tilde{d}$ and $\tilde{l}$ to denote the consumption of nonmedical goods and leisure under the ethical

option. These two variables are derived by substituting the solution for $\tilde{s}$ into expressions (3) and (2) respectively.

Now we can calculate the critical level of ethical costs, denoted as e^*, for which the physician is indifferent between the financial and the ethical option by solving the equality:

$$\begin{aligned} v_k(\bar{d}_k, \bar{l}_k, e^*) &= v_k(\tilde{d}_k, \tilde{l}_k, 0) \quad \Rightarrow \\ e^* &= \bar{v} - \tilde{v} \\ &= \left(\alpha \bar{d}^{-\rho} + \bar{l}^{-\rho}\right)^{-1/\rho} - \left(\alpha \tilde{d}^{-\rho} + \tilde{l}^{-\rho}\right)^{-1/\rho} \end{aligned} \tag{6}$$

where $\bar{v}$ and $\tilde{v}$ are implicitly defined by the first and last term of the second line of equation (6) respectively.

Using this expression for e^*, the outcome of the optimization problem can now be summarized as follows:

$$\begin{aligned} s_k &= \bar{s} \qquad e_k \le e^* \\ s_k &= \tilde{s} \qquad e_k \ge e^* \end{aligned} \tag{7}$$

A physician thus prefers the financial option if e^* exceeds her ethical cost variable e. The ethical option is chosen whenever e^* is below this threshold value. By definition, the physician is indifferent between the two options when $e = e^*$.

We now assume that physicians are heterogenous with respect to the value of the ethical cost variable e. Hence, a number of physicians will prefer the ethical option while others will choose the financial option. Denoting the distribution function of e by $G_s(e)$, the supply of medical services at the aggregate level can then be expressed as a weighted average of supply as defined for the two options:

$$S = G_s(e^*)N_s\bar{s} + (1 - G_s(e^*))Z \quad (8)$$

How does medical supply react to fee changes? The supply by ethical specialists is not affected by the services fee as it equals demand and all patients are fully insured (see below). In case of the financial option, we distinguish two direct effects that take the opposite direction. A higher fee makes it financially more attractive to provide medical services (a substitution effect), but also increases the physician's income and thereby her demand for leisure (an income effect). In the numerical version of our model, the substitution effect dominates. One may argue that as income effects of fee changes are neutralized by adjustments in the lumpsum income the income effect is nil. However, this is not entirely correct. The point is that the lumpsum income adjustment is the same for the ethical and financial specialists, whereas supply and thus the income effects are different for both groups. As income is relevant only for the supply behavior of financial specialists, redistribution from ethical towards financial specialists may reduce aggregate supply.

An important point to note is that low fee values may restrict the supply by financial specialists to be zero as supply cannot take a negative value. Therefore we expect the supply of services by financial specialists to be a function of the services fee that consists of two parts: a flat part for low fee levels and an upward-sloping part for high fee levels.

Given that the supply by ethical specialists is unresponsive to the fee for services, the aggregate supply function should have properties identical to the function that describes the supply by financial specialists.

However, from equation (6) it follows that a composition effect may also play a role as total supply also depends on the share of specialists $G_s(e^*)$ that prefers the financial option.

Intuitively, it will be clear that the fraction of financial specialists is a positive function of the service fee in deviation from the fee level that equalizes supply and demand. At this equilibrium level, all specialists prefer the ethical option as both options yield the same levels of consumption and leisure, but only the financial option includes ethical costs. If the fee deviates from this specific level, the financial option generally corresponds to a different pair of consumption and leisure and a higher level of utility. Hence, the change in the fee away from the equilibrium level drives a wedge between the consumption-leisure pairs that correspond to the ethical and financial option and thus widens the utility gap between both options. Consequently a number of specialists, in particular those with the lowest ethical costs, will switch to the financial option. This argument works also the other way round. Hence, we expect the fraction of financial specialists to be U-shaped function of the fee, with a minimum of zero at the fee level that equalizes demand and supply.

To sum up: what is the impact of an increase in the service fee on medical supply? The aggregate direct effects pushes up supply. The direction of the indirect or compositional effect depends on whether we move towards or away from the equalizing fee level and which of the two options features the highest supply.

Patients

The patient derives utility from the consumption of medical specialist services, z, and the consumption of other goods and services, c. Patients are assumed to have linear-quadratic additively separable preferences over the two types of goods,

$$u_i = c_i - \frac{1}{2}\gamma c_i^2 + \varepsilon_i z_i - \frac{1}{2}\delta z_i^2$$
$$\delta > 0,\ \varepsilon_i \geq 0,\ 0<\gamma<1/y \tag{9}$$

where u_i is patient i's utility and y is the income of the patient net of health care premiums.

Rather than using equation (9), we simplify by imposing full insurance. The effect of full insurance is that nonmedical consumption equals the patient income net of health care premiums (given our nonsatiation assumption in (9)): $c_i = y$. As will become clear below, income net of health care premiums is not affected by the reform. Hence, the first two terms at the RHS of (9) reduce to a constant and maximization of u_i in (9) is equivalent with maximizing u'_i in (10):

$$u_i' = \varepsilon_i z_i - \frac{1}{2}\delta z_i^2 \tag{10}$$

Our motivation to employ a linear-quadratic specification is twofold. First, we want to allow the marginal utility of medical care to turn negative if medical consumption becomes sufficiently high (*e.g.* due to time costs or iatrogenesis). Second, zero marginal utility of medical care is needed to obtain a finite solution for demand in case of a zero out-of-pocket price.

The solution to the patient's optimization problem reads as follows:

$$z_i = \frac{\epsilon_i}{\delta} \tag{11}$$

We allow medical need to be unevenly distributed among individuals. This heterogeneity is reflected in the model by assuming ϵ to differ across individuals. Demand for medical services at the population level can then be expressed as follows, using N to denote the population of patients:

$$Z = N\frac{\mathrm{E}(\varepsilon_i)}{\delta} \tag{12}$$

Closure of the model

We close the model by specifying the calculation of premiums, the consumer price of medical services and the lumpsum subsidies to specialists.

Health care premiums, P, equal aggregate health expenditure, $t_c S$. The consumer price of medical services, t_c, consists of the fee-for-service, t, plus a tax levied at rate τ:

$$t_c = t(1+\tau) \tag{13}$$

The revenues from this tax, τtS, are used to finance lumpsum subsidies to specialists, $N_s h$:

$$\tau = \frac{N_s h}{tS} \tag{14}$$

By aggregation of equation (3) over all specialists we obtain

$$D = N_s h + tS \tag{15}$$

where $D = N_s \int d_k dG_s(e_k)$ and S is defined in equation (8). This equation can be written as an expression for h:

$$h = \frac{D - tS}{N_s} \tag{16}$$

This expression tells us which value of h keeps D constant when the government changes the fee level t.

The neutrality of physician income with respect to the fee for medical services implies that health care expenditure is a constant. The same holds true for health care premiums and for nonmedical consumption of patients, *c,* which equals patient income net of health care premiums. The welfare analysis will make use of this property.

Welfare

Individual patient welfare can be derived basically by substituting the solution for medical consumption into the patient's utility function. Note that this solution for consumption will in general deviate from demand. As a rationing rule, we assume that the specialist who prefers the financial option divides her services between all her patients so that each patient faces the same discrepancy between demand and supply in relative terms (proportional rationing).

Total patient welfare U is defined as the sum of individual utilities, where the summation runs over consumers who differ in their value for ϵ_i:

$$U = N\int_0^{\infty} u_i' \, dG(\varepsilon_i) = N\int_0^{\infty} \left[\varepsilon_i s_i - \tfrac{1}{2}\delta s_i^2\right] dG(\varepsilon_i) \tag{17}$$

Similarly, welfare of medical specialists, V, is defined as the sum of the welfare of specialists choosing for the ethical and the financial option:

$$\begin{aligned} V &= \bar{V} + \tilde{V} - N_s E(e_k \mid e_k < e^*) \\ &= N_s\left[(1 - G_s(e^*))\bar{v} + G_s(e^*)\tilde{v} - \int_0^{e^*} e_k \, dG_s(e_k)\right] \end{aligned} \tag{18}$$

where $\bar{V}$ and $\tilde{V}$ are implicitly defined by the first and second term on the second line of equation (18).

In equation (18), $G_s(e^*)$ equals the share of specialists that opt for the financial option. Ethical costs are only relevant for those specialists who choose to bear these costs, *i.e.* the specialists whose e_k is below e^*.

We also want to calculate social welfare. Note that simply adding up patient welfare (equation (17)) and specialist welfare (equation (18)) makes no sense as the two welfare functions are non-comparable. Therefore, we first transform the changes in patient and physician welfare into their consumption-equivalent counterparts. The latter are defined in terms of the nonmedical consumption good which is consumed by both patients and specialists. The consumption-equivalent change in social welfare is then the sum of the consumption-equivalent changes of patient welfare and physician welfare. Our discussion paper gives a detailed specification of the method of consumption-equivalent welfare changes.

Note that our measure of social welfare neglects redistributional concerns. The reason is that the focus of our analysis is more on efficiency effects than on distribution effects (witness also our assumption of proportional rationing).

EVALUATION OF POLICY CHANGES

The numerical model

Actually, the numerical model that we use for our calculations includes more institutional detail than the model set up so far. In particular, it adds three elements. First, it decomposes the patient population into a group that is covered by public insurance and a group that is privately insured. This allows us to take into account

differences in the average income levels of both groups and in the distribution of health care needs. Second, the privately insured face a copayment scheme that includes deductibles. Consequently, their demand also depends on the (out-of-pocket) price of medical services and of income net of health care premiums (see [20] for details). Finally, the numerical model distinguishes between first and subsequent treatments. The former are assumed to be demand determined; the consumption of subsequent treatments is determined by physicians in the way that has been discussed in the previous section.

These elements add to the realism of the model. At the same time, they make only a small quantitative contribution to our simulations, justifying that we left them out in the discussion of our model in the previous section. In addition, we should note that, before the financial reform, fees for medical specialist services delivered to publicly insured patients were lower than the fees of services to the privately insured. However, since we want to focus exclusively on the move from a fee-for-service to a capitation scheme, the simulations we present impose that the two fees are equal also prior to the reform.

In addition, the numerical model chooses values for all parameters and exogenous variables in the model. Parameter values are found by a combination of estimation and calibration. Details can be found in the working paper version of this paper ([20]).

A base simulation

We evaluate the values of interesting variables at 13 levels for the fee for services, which we vary from 0 to euro 1080 in steps of euro 90. This seems to be too general for the purpose at hand; for an evaluation of the financial reform, we only need to compare outcomes before and after the reform (the average fees equal t=euro 360 and t=0 respectively). However, by evaluating a wider range of fee values, we get a better picture of what is going on between t=0 and t=euro 360.

Before the reform, lumpsum subsidies to medical specialists are zero. After the reform, the fee takes a value of zero so that the income of specialists consists entirely of lumpsum subsidies. At fee levels higher than euro 360, lumpsum subsidies to physicians are negative. The revenues of these negative subsidies are reimbursed to patients by letting the consumer price be below the fee for services.

We first analyze the supply curves of physicians that prefer the ethical or financial option. As Figure 1 shows, the supply curve of the ethical specialist is (almost) flat. As ethical specialists follow demand, there are no direct effects of changes in the fee for services, only indirect effects which operate through the demand for specialist services (the supply curve for ethical specialists reflects the aggregate demand curve).

The supply curve of financial specialists is more complex. When the fee is below euro 270, the zero corner solution applies. When the fee is increased beyond the value of euro 270, supply becomes positive and increasing. Hence, the substitution effects of fee changes dominate the income effects. Due to the curvature of the utility function of the specialist, the elasticity of labor supply with respect to the fee for services falls when the fee is further increased.

insert figure 1 about here

Supply under the ethical and under the financial option are equal at a fee value t^* of about euro 630. Hence, the utilities for the two options (exclusive ethical costs) also coincide. The consequence of this is illustrated in Figure 2 which shows the share of physicians who prefer the financial option. At $t = t^*$ = euro 630 this share equals zero. The more the fee t deviates from t^*, the higher the yield of optimization and the larger the number of medical specialists who prefer the financial option.

Figure 1 demonstrates that the aggregate supply of subsequent treatments is increasing everywhere. That the slope of the aggregate supply curve is positive for low values of the fee ($t <$ euro 270), follows entirely from the negative relationship between the number of financial medical specialists and the fee (see Figure 2). The decline in the number of financial physicians increases the weight of ethical specialists who supply more services. For $t >$ euro 270, the positive slope of the aggregate supply curve must also be attributed to the upward-sloping supply curve of financial physicians.

insert figure 2 and figure 3 about here

Figure 3 shows the income of specialists. By construction, average income of financial and ethical specialists is independent of the fee for services. However, the same does not hold true for financial and ethical specialists taken separately. The reason is that their supply values are different whereas lumpsum subsidies are equal.

Because a fee increase boosts the supply of services, it reduces lumpsum subsidies to specialists. As labor supply in the financial option equals zero for small fees, it thus decreases the income of financial specialists. Similarly, income of ethical specialists increases. This also explains why fee increases reduce the number of specialists that prefer the financial option when the fee is small (see Figure 2). For higher fees, the reverse holds true. For $t >$ euro 270, the labor supply of the financial specialist is increasing in the fee for services, whereas demand is almost independent of this fee. Hence, the income of the financial specialist is increasing and the income of the ethical specialist is decreasing in the fee for services.

However, the case where t moves from euro 270 to euro 360 is an exception to this rule. Both the incomes of financial physicians and of ethical physicians decrease. This can only be consistent with constancy at the macro level if there is a compositional effect. Indeed, there is: upon the increase of the fee for services from t=270 to t=360, the number of financial medical specialists falls. Since ethical specialists produce more than financial specialists at these fee levels, this increases medical production and makes it necessary for the government to economize on lumpsum subsidies (see equation(16)). This explains why both types of physicians face a fall in income.

insert figure 4 , 5 and 6 about here

Figures 4, 5 and 6 show the consumption-equivalent changes in patient welfare, specialist welfare and social welfare respectively. Recall that consumption of non-medical services is invariant to fee movements. Hence,

changes in patient welfare fully reflect movements in health care consumption. Figure 4 shows that patient welfare reaches its peak value at $t = t^*$, where supply equals demand.

How can we explain that patients face the highest welfare gain when they move to a situation in which supply meets demand? Obviously, we would expect differences between supply and demand to be welfare-reducing, as the patient chooses his demand such as to maximize his utility. So in case supply meets demand, welfare losses due to imperfect agency are zero. However, the moral hazard argument suggests that welfare could be raised by moving supply below demand. The reason why this does not occur lies in our financing rule for specialist income. The welfare gain from a smaller loss due to moral hazard is reflected in lower health insurance premiums and lower copayments. However, our simulations keep the sum of health insurance premiums and copayments equal to the budget value. The latter is kept constant by adjusting the lumpsum income to compensate for fee changes. Hence, any welfare gain from the reduction of moral hazard is transmitted to the group of medical specialists. Therefore, the welfare gain of the patient must be optimal if in the new situation supply meets demand.

Figure 5 demonstrates how the consumption equivalent change in specialist welfare depends on the fee value. The maximum is achieved by moving to a situation in which the fee equals zero. This is well below the fee that optimizes patient welfare, reflecting the redistribution to physicians of the welfare gains that the elimination of moral hazard brings about. Note that if we move from the initial situation (t=360) to a zero fee value the number of specialists that prefer the financial option increases, whereas if we move from t=360 to higher fee values the number of ethical specialists increases (see figure 2). Hence the group of specialists that

switches between both options consists of ethical specialists that prefer the financial option in the new situation (in case the fee value is between 0 and 360) and financial specialist that choose to be ethical after the reform (fee values higher than 360). Hence, the consumption-equivalent change of the welfare of the switching group should be higher than the change for ethical specialists if the fee value is below 360 and higher than the change for financial specialists when the fee is above the 360 value; otherwise there would be reason to change.

Maximum social welfare is achieved at a fee of euro 540 (see Figure 6). This value is thus below the level for which supply equals demand. Optimal supply policies reduce supply in order to compensate for the moral hazard effect in the demand for physician services.

What was the impact of the financial reform? Economically, the reform induced physicians to reduce the supply of medical services. First, physicians who prior to the reform already preferred the financial option reduced their supply of services. Second, physicians with relatively low ethical costs switched to the financial option, thereby also reducing their supply.

Apart from this a general equilibrium effect reinforced the effect upon supply. With lower supply, the government had to raise lumpsum subsidies to physicians in order to keep their incomes at their original levels. Therefore, despite the removal of the fee for services, the consumer price went up. This depressed demand and thereby contributed to the reduction of supply. However, this general equilibrium effect is quite small quantitatively, as witnessed by the graphs.

Patients faced a decline in welfare - due to the fall of medical consumption. In contrast, physicians benefitted from the increased amount of leisure time. Social welfare decreased, as demonstrated in Figure 6. In summary, the policy reform eliminated the moral hazard in medical consumption but it also aggravated the distortion due to imperfect agency. On net, the reform was welfare-reducing.

Sensitivity analysis

The conclusion of the previous section may be sensitive to the values of model parameters. Therefore, this section focuses on two variables that we view as crucial for our numerical results: the price elasticity of supply and the average income of medical specialists. Therefore, in this section we discuss two experiments. In the first we vary the price elasticity of labor supply by imposing 75% higher and lower values for the elasticity of substitution σ between leisure and consumption in the utility function of the specialist. Then we examine the consequences of having the budget of physicians linked to fee values of euro 180 and euro 540, respectively.

We have deliberately specified the range of parameter values defined by these calculations as very wide and even wider than we judge as realistic. The reason is that if our conclusions pass this test, this strongly supports the results from the benchmark simulation.

In the base simulation, the elasticity of substitution σ is set at 6.91. Here we analyze effects of changing its value to 3.9 (low variant) and 12.1 (high variant). In the high variant, the fee value that minimizes the number of financial specialists falls down from euro 630 to about euro 450. The same holds true for the fee that

optimizes the patient welfare function. In the low variant ($\sigma = 3.9$) it appears that the share of financial specialists does not reach its minimum at fee values below euro 1080. We observe the same phenomenon when we inspect social welfare. In this case, the model is always in a situation of excess demand. This is linked to the strong preference of medical specialists for leisure time relative to consumption. We conclude that our results are quite sensitive to changes in the price elasticity of supply of medical specialists. However, our main conclusion that the introduction of local initiatives was welfare-reducing remains valid in the two alternative cases.

A second parameter to explore in the sensitivity analysis is the value of the specialist budget. The average budget per specialist in the base scenario amounts to euro 129,200, which corresponds to the average income in absence of a budget when the fee equals euro 360. In the low scenario the budget is reduced to euro 71,700; in the high variant the budget is euro 219,500. These income values are linked to situations where the budget is zero and the fee equals 180 and 540 euros, respectively.

The minimum of the fraction of specialists who prefer the financial option varies from fee values of (about) euro 540 (low variant) to euro 720 (high variant). The optimal fee values differ only slightly more than in the two previous cases: euro 630 in case of a high budget and euro 450 in case of a low budget. Again, our main result that the shift in financing arrangement for medical specialists was welfare-reducing continues to hold true.

Table 1 summarizes our findings. It also includes the results of changing the price elasticity of health care demand by patients. In [20] we also varied the price elasticity of demand through 75% upward and downward

changes in the parameter γ in the utility function (9). The results of the model were quite insensitive to these variations and the financial reform still appeared to be welfare reducing.

insert table 1 about here

CONCLUSIONS

This paper draws two conclusions. First, the reduction of fees for specialist services in the Netherlands reduced social welfare. This result follows from the position of the initial fee for services, euro 360, relative to the level that optimizes our social welfare measure, euro 540. As the policy change increased the deviation of the actual fee from its optimal level, it aggravated the welfare loss from a suboptimal financing scheme.

Note that this conclusion does not hinge upon the position of the optimal fee. As our sensitivity analysis and Figure 6 clearly demonstrate, we would have found the same result when the optimal fee level had been somewhat below euro 360. Indeed, the steep decline of the social welfare measure in the range running from a zero fee to a fee of euro 360 indicates that our result on the welfare-decreasing nature of the shift in financing scheme is robust to minor changes of parameters or initial conditions.

The conclusion does not depend on the exact definition of social welfare either. One may argue that social welfare should be defined exclusive of physicians. Figure 6 clearly demonstrates that such a change of definition would leave unaltered our central finding that the reform imposed a welfare loss.

Neither does our conclusion depend upon the definition of the policy change. We have defined the financing scheme that resulted after the reform as being fully lumpsum, *i.e.* with a zero services fee. Van den Berg and Mot [18] report that in some hospitals, it is agreed that specialist budgets might be increased if specialist production grows faster than expected. To the extent that these agreements are credible, the financial reform may actually have been perceived not as a switch from a fee-for-service scheme to a lumpsum scheme, but as a plain reduction of the fee for medical services. It can be seen from Figure 6 that, if this is true, our calculations overstate the impact of the change in financing scheme. However, it would not affect our finding that the introduction of a lumpsum budget for specialist services has reduced social welfare.

A second result of our analysis is that the optimal financing scheme features both a fee-for-service element and a lumpsum element. In particular, in order to be able to meet two targets, *i.e.* an efficient supply of medical services and a target income for physicians, one should have at least two instruments. A financing scheme that combines a fee-for-service element with a lumpsum element provides two such instruments. Restricting the financing scheme to one instrument only will generally result in a suboptimal outcome.

Our analysis could be extended further. One option is to explore whether the financial reform has induced specialists to engage in risk-selection strategies. Indeed, fee reductions may have induced medical specialists to intensify their efforts to dump high-risk consumers, *i.e.* supply no services at all to these patients ([21-22]). If this is the case, the financial reform may have had more adverse welfare effects than recognized in this paper.

REFERENCES

1. Rosett, R.N. and Huang, L. The Effect of Health Insurance on the Demand for Medical Care. *Journal of Political Economy* 1973; **81**: 281-305.

2. Newhouse, J.P. and the Insurance Experiment Group. *Free for All?, Lessons from the RAND Health Insurance Experiment*. Cambridge: Harvard University Press, 1993.

3. Evans, R.G. Supplier-induced Demand: Some Emperical Evidence and Implications. In: Perlman, M., editor. *The Economics of Medical Care*. London: MacMillan, 1974: 162-173.

4. Fuchs, V.R. (1978), The Supply of Surgeons and the Demand for Surgical Operations, *Journal of Human Resources* 13, Supplement, 35-56.

5. Newhouse, J.P. (1987), Medical Care Expenditure: A Cross-National Survey, *Journal of Health Economics* 6, 109-128.

6. Folmer, C., Stevens, J., Tulder, F. van, and Westerhout, E.W.M.T. Towards an Economic Model of the Dutch Health Care Sector. *Health Economics* 1997; **6**: 351-363.

7. Canton, E.J.F. and Westerhout, E.W.M.T. A Model for the Dutch Pharmaceutical Market. *Health Economics* 1999; **8**: 391-402.

8. Feldman, R. and Dowd, B. A New Estimate of the Welfare Loss of Excess Health Insurance. *American Economic Review* 1991; **81**: 297-301.

9. Manning, W.G. and Marquis, M.S. Health Insurance: the Tradeoff between Risk Pooling and Moral Hazard. *Journal of Health Economics* 1996; **15**: 609-639.

10. Zeckhauser, R. Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives. *Journal of Economic Theory* 1970; **2**: 10-26.

11. McGuire, Th.G. and Pauly, M.V. Physician Responses to Fee Changes with Multiple Payers. *Journal of Health Economics* 1991; **10**: 385-410.

12. Rizzo, J.A. and Blumenthal, D. Physician Labor Supply: Do Income Effects Matter?. *Journal of Health Economics* 1994; **13**: 433-453.

13. Wedig, G., Mitchell, J.B. and Cromwell, J. Can Price Controls Induce Optimal Physician Behavior?. *Journal of Health Politics, Policy and Law* 1989; **14**: 601-620.

14. Ellis, R.P. and McGuire, Th.G. Optimal Payment Systems for Health Services. *Journal of Health Economics* 1990; **9**: 375-396.

15. Ellis, R.P. and McGuire, Th.G. Supply-Side and Demand-Side Cost Sharing in Health Care. *Journal of Economic Perspectives* 1993; **7**: 135-151.

16. Mot, E.S. (2001), The Influence of the Payment System for Medical Specialists upon Waiting Times in the Netherlands, mimeo.

17. Ziekenfondsraad (1998), *Evaluatie experiment specialistenhonorering*, rapport no. 783, Amstelveen (in Dutch).

18. Berg, B. van den, and Mot, E.S. Honorering, verwijsgedrag en prikkels. *Economisch Statistische Berichten* 2000; **85**: 791-793 (in Dutch).

19. Statistics Netherlands (2000), Vademecum of Health Statistics of the Netherlands, Voorburg .

20. Folmer, C. and Westerhout, E.W.M.T. (2002), Financing Medical Specialist Services in the Netherlands: Welfare Implications of Imperfect Agency, CPB Discussion Paper no. 6, The Hague.

21. Ellis, R.P. and McGuire, Th.G. Provider Behavior under Prospective Reimbursement - Cost Sharing and Supply. *Journal of Health Economics* 1986; **5**: 129-151.

22. Newhouse, J.P. Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection. *Journal of Economic Literature* 1996; **34**: 1236-1263.

Figure 1 Supply of treatments

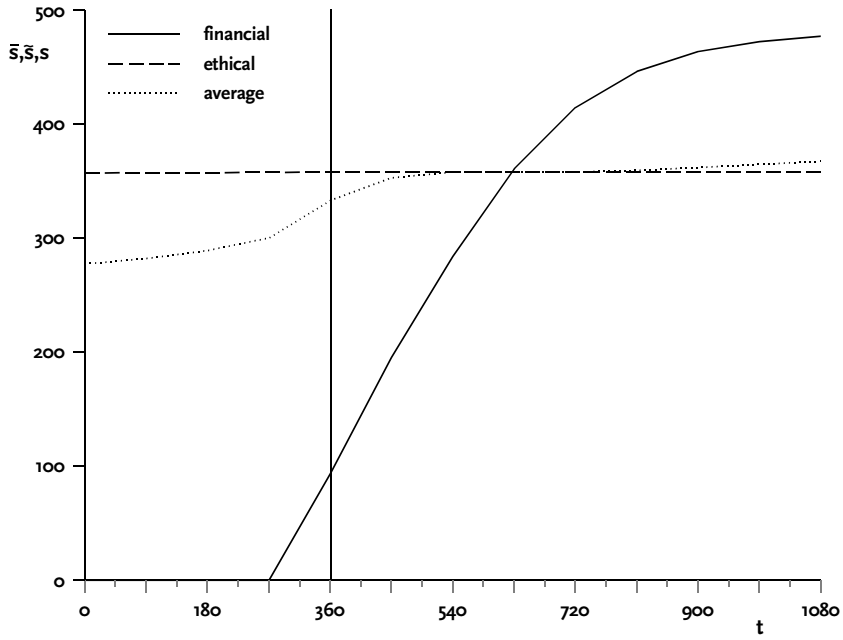

Figure 2 Share of specialists that prefer the financial option

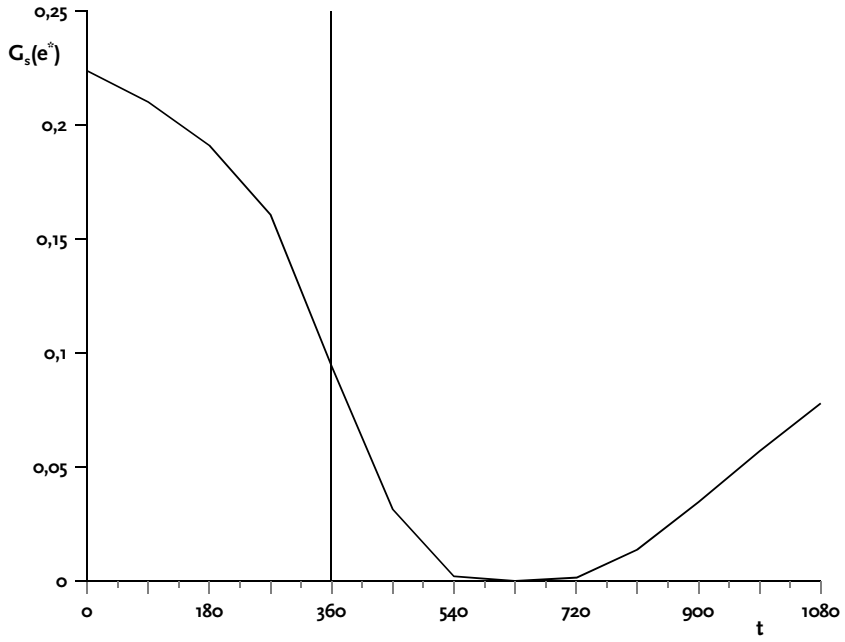

Figure 3 Income per specialist

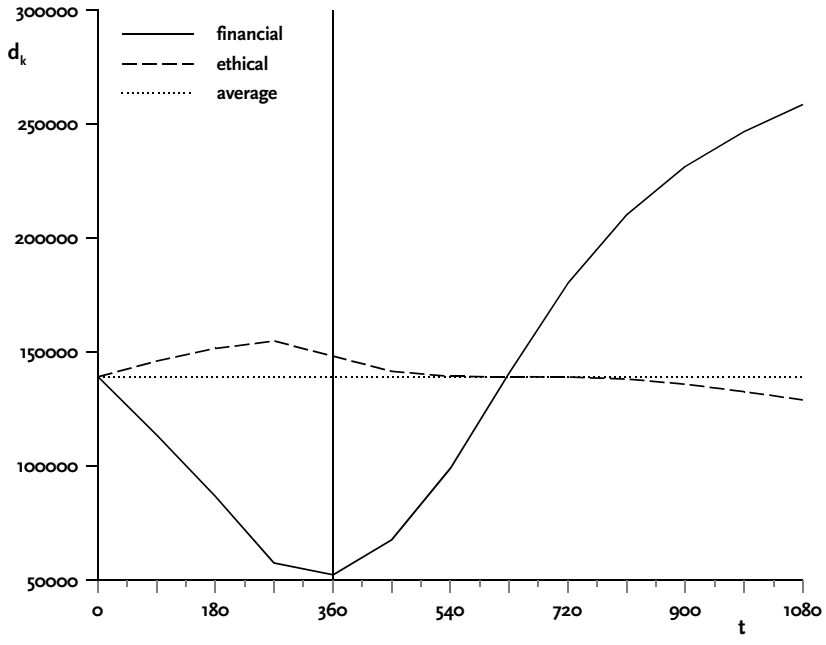

Figure 4 Consumption-equivalent change of patient welfare

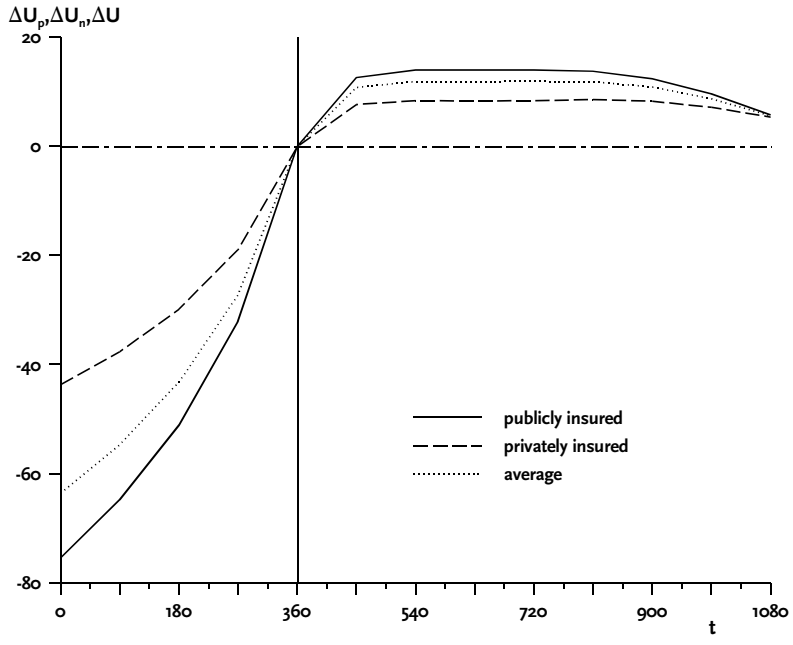

Figure 5 Consumption-equivalent change of medical specialist welfare

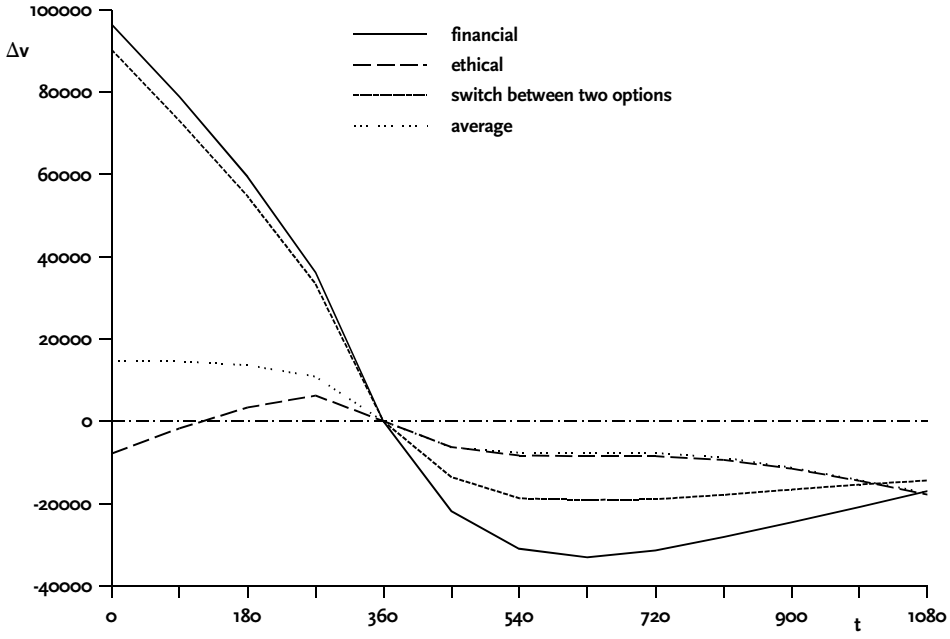

Figure 6 Consumption- equivalent welfare changes

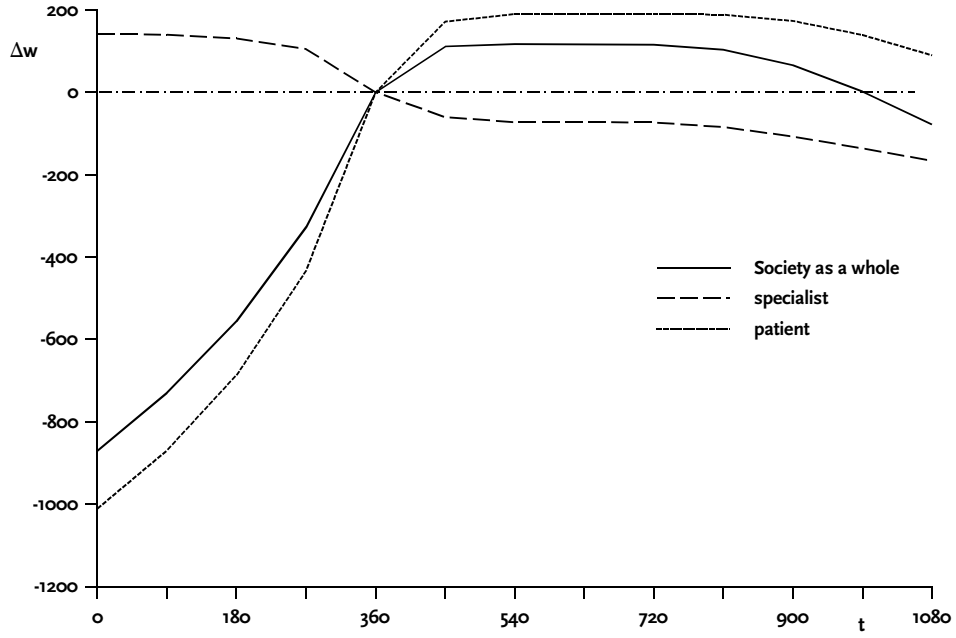

Table 1 Results corresponding to 4 alternative parameter configurations

variant	optimal fee (patients)	optimal fee (specialists)	optimal fee (total)	welfare change from old to new
benchmark	540	0	540	–870
high σ	450	0	450	–980
low σ	1080	180	>1080	–345
high budget	810	0	630	–932
low budget	450	0	450	–848