CPB Discussion Paper

No 78

February 2007

Co-payment systems in health care: between moral hazard and risk reduction

Ed Westerhout and Kees Folmer

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone	+31 70 338 33 80
Telefax	+31 70 338 33 50
Internet	www.cpb.nl

ISBN 90-5833-311-7

Contents

Abstract in English 5

Abstract in Dutch 5

1 Introduction 7

2 A model of health care consumption 11

3 A simplified version 15

4 A more realistic model 19
4.1 Optimization 19
4.2 Optimal co-payment policies 22

5 Calibration of model parameters 27

6 Numerical calculations 31

7 Sensitivity analysis 37

8 Concluding remarks 39

Appendix A Critical values for the health care need parameter 41

Appendix B Moments of the lognormal distribution 43

References 45

Abstract in English

It is well-known that co-payments in health insurance may increase social welfare by reducing moral hazard. Considerably less is known about the form co-payment schemes should ideally take. This paper investigates what co-payment rate and co-payment maximum characterize the optimal scheme, *i.e.* the scheme that achieves the highest level of social welfare, within the class of two-part co-payment schemes of which the second part features a zero rate. It also quantifies the welfare losses that correspond with sub-optimal co-payment schemes.

The paper uses a model with optimizing households that are risk-averse, exercise price-elastic demand and are aware of the kinks in their budget constraints. Numerical simulations with this model indicate that the optimal scheme combines a 80% rate with a maximum of about 600 euro. Sensitivity analysis shows that the maximum varies a lot with changes in basic parameters; the 80% value for the optimal co-payment rate is quite robust, though. The welfare losses that correspond to alternative co-payment schemes are generally quite small.

Abstract in Dutch

Het is bekend dat eigen betalingen in zorgverzekeringen welvaartsverhogend kunnen zijn omdat ze het oneigenlijk gebruik van zorgvoorzieningen beperken (moral hazard). Over de ideale vormgeving van eigen betalingen bestaat veel minder duidelijkheid. Dit onderzoek richt zich op de vormgeving van het optimale systeem van eigen betalingen binnen de groep van systemen die zich karakteriseren door één bijbetalingsvoet en bijbetalingsmaximum. Ook worden de welvaartsverliezen van niet-optimale systemen gekwantificeerd.

Het onderzoek gebruikt een model met optimaliserende huishoudens die risico-avers zijn, prijsgevoelig en bewust van de vorm van hun budgetrestrictie. Numerieke simulaties met het model geven aan dat het optimale systeem van eigen betalingen een voet van 80% combineert met een maximum van 600 euro per verzekerde. Het maximum is vrij gevoelig voor keuzes van modelparameters; de optimale bijbetalingsvoet is echter vrij robuust. Welvaartsverliezen van niet-optimale systemen zijn relatief klein.

Keywords: Moral Hazard, Deductibles, Co-payments
JEL codes: D60, H21, I18

1 Introduction[1]

Health insurance schemes are widespread. The unpredictable nature of medical consumption in a number of respects (timing, frequency, intensity, costs) makes it unthinkable that medical consumption would go uninsured (Arrow (1963)). Ideally, health insurance would take the form of payments that are only conditional on the health status of the insured (Blomqvist (1997)). Then, under certain conditions the first-best solution would be attainable. However, information on the health status of the insured is private, if it is available at all. This explains why health insurance typically takes the form of payments that are conditional on health expenditure. As a result, health insurance implies a moral hazard distortion[2] and partial insurance will generally yield higher welfare than full insurance (Pauly (1968)).

Indeed, many insurance schemes use co-payment elements. Often, co-payments take the form of proportional schemes, schemes which adopt a co-payment rate smaller than one (Robinson (2002)). Deductible schemes can also be found in a number of countries, in particular in the private insurance sector. But there are many other forms of co-payment schemes. Further, the classes of proportional schemes and deductible schemes are very heterogeneous, in terms of maxima and in terms of services for which co-payment is required, for example.

Ideally, co-payment schemes are constructed such that they strike a balance between risk sharing and moral hazard. Hence, it is important to know what form co-payment schemes should ideally take. In particular, should we have co-payment rates somewhere between zero and one that are independent of the neediness of the insured? Should insurance schemes apply a co-payment maximum beyond which no further co-payments are required? Would it be optimal in general to apply more than one co-payment rate? Should co-payment rates be increasing or decreasing in the amount of health care spending? Quite surprisingly, though, these issues have rarely been discussed in the literature.

This paper characterizes the form of the optimal co-payment scheme in the class of two-part co-payment schemes of which the second-part co-payment rate is zero. It adopts a social welfare function that is derived explicitly from the utility functions of households that are risk-averse, price-responsive and aware of the nonlinearity of their budget constraints. Because of the nonlinearities in the model, we are unable to derive closed-form solutions. Therefore, we follow two approaches. First, we simplify the model to derive results on the basis of its analytical solution. Second, we solve the full model numerically. The numerical results reiterate

[1] An earlier version of this paper has been presented at the 1997 European Workshop on Econometrics and Health Economics, at the 1997 NAKE Research day and at the 2005 IIPF Conference. Thanks are due to the participants at those meetings, to our discussants Kris de Jaegher and Erling Holmoy. Rob Waaijers processed individual income data to and generated a distribution of patient income. Further, we thank Eddy van Doorslaer, Bas Jacobs and three referees for useful suggestions. Any remaining errors are the responsibility of the authors.

[2] Actually, there are several types of moral hazard (see Zweifel and Manning (2000) for an overview). We restrict ourselves to the ex post type of moral hazard, which is the type that is most discussed in the literature.

the results from the analytical approach, but add an indication of the numerical configuration of the optimal co-payment scheme, of the differences across co-payment schemes and of the welfare losses that correspond to sub-optimal schemes.

There is some earlier literature on the optimal form of co-payment schemes. Arrow (1963) and Raviv (1979) discussed the optimality of deductibles, but neglected the price-responsiveness of health services demand. Feldstein (1973) and Feldman and Dowd (1991) did account for moral hazard considerations, but assumed that the budget constraints of households are linear. Keeler *et al.* (1977) and Ellis (1986) explored health care demand in case of a nonlinear budget constraint. The former of these two articles touched upon the optimality of different co-payment schemes. Unlike our analysis however, it did not connect the social welfare function with the utility function from which health care demand is derived. This renders it difficult to assess on a consistent basis the welfare implications of different co-payment schemes.

The analysis by Manning and Marquis (1996) made an improvement precisely on this point. Their analysis of the optimal trade-off between risk pooling and moral hazard is based on estimates of risk aversion and the price elasticity of demand, which ultimately are based on one model of household behaviour - a model that explains both the demand for health care and that for health insurance. However, they were unable to find plausible estimates of the optimal stop loss with the data they used. According to the authors, this may be due to rather imprecise estimates of the risk aversion of households.

Blomqvist (1997) characterizes on a general level the form of the optimal non-linear co-payment scheme. His analysis applies the theory of optimal taxation along the lines of Mirrlees (1971) to the problem of optimal health insurance without imposing any restrictions on the form of this scheme. Our approach is more limited as we restrict the analysis to schemes with two linear segments of which the second segment features a zero rate. The reason for doing so is that in reality co-payment schemes often feature a small number of segments only, resembling more our two-part scheme than Blomqvist's continuous scheme.[3]

The structure of our paper is as follows. Section 2 sets up a general model of health care demand with a kinked budget constraint. Section 3 presents a simple version of this model in which medical need obeys a two-spike distribution function. This version specifies a Ramsey rule for the optimal co-payment scheme and finds that deductibles can never be optimal. Section 4 adopts a more detailed version of the model using a more realistic distribution of medical need. An analytical solution can then no longer be obtained and we have to rely on numerical simulations. Section 5 describes how we have calibrated this version of the model on data for the Netherlands. The simulations in section 6 reiterate the result suggested by the Ramsey rule: within the class of co-payments schemes we consider a deductible scheme is

[3] It is an interesting question how to explain this phenomenon. One possible explanation may relate to administration costs. The costs of running a co-payment scheme may be strongly increasing in the complexity of the scheme.

suboptimal. We calculate that the co-payment rate that corresponds with the optimal co-payment scheme is in the order of 80%. Section 7 explores the sensitivity of our conclusions to changes in an number of important parameters. Section 8 contains concluding remarks.

2 A model of health care consumption

The representative patient[4] has quadratic preferences for the consumption of health care, denoted z, and that of other, non-medical services, denoted c. Using u to denote direct utility, the consumer's direct utility function can be written as follows:

$$u = c - \tfrac{1}{2}\varepsilon_c c^2 + \varepsilon_z z - \tfrac{1}{2}\varepsilon_m z^2$$
$$\varepsilon_c,\ \varepsilon_m > 0,\ \ \varepsilon_z \geq 0 \tag{2.1}$$

We postulate the quadratic form for the relation between utility and medical consumption mainly for two reasons. First, to ensure that the demand for medical services is always finite, *i.e.* also in the case of a zero out-of-pocket price[5]. Second, for analytical tractability. The quadratic form allows us to express expected utility as a function of the first and second moments of variables. This allows us to derive the optimal co-payment structure by means of calculation rather than stochastic simulation.

In addition, the quadratic form has some attractive properties. It implies that the price elasticity of health care demand decreases if the health status of the patient worsens. Such a relationship is not only intuitive, it is also backed by empirical evidence (Wedig (1988)). Furthermore, this price elasticity is increasing in the co-payment rate, again a feature for which there is empirical evidence (Rosett and Huang (1973), Phelps and Newhouse (1974), Newhouse and the Insurance Experiment Group (1993)).

Equation (2.1) shows that we adopt the quadratic form also for the relation between utility and the consumption of non-medical services, but to avoid that the marginal utility of non-medical products becomes non-positive, we impose that ε_c is not "too high"[6].

The parameter ε_z differs between patients. This reflects patient heterogeneity in terms of their need of health care (the marginal utility of medical consumption equals $\varepsilon_z - \varepsilon_m z$). Underlying is a state-dependent health production function in which the value of medical care is a decreasing function of the health status of the patient: the worse the health of a patient, the more beneficial will be medical intervention. The other parameters in equation (2.1) are the same for all consumers. For brevity, we do not index ε_z explicitly.

The corresponding budget constraint is piecewise linear. We distinguish two linear pieces:

[4] Throughout the paper, we will talk about patients for simplicity, although, as we will see below, it may be the case that a patient decides not to consume any medical care.

[5] The standard utility function with positive marginal utility everywhere does not yield an interior solution in case the out-of-pocket price is zero.

[6] With the budget constraint given below, it can be derived that this implies that $\varepsilon_c < 1/(y\text{-}p)$.

$$c = y - p - b_1 tz \qquad 0 \le z \le m/(b_1 t)$$
$$c = y - p - m - b_2 t\left(z - \frac{m}{b_1 t}\right) \qquad z \ge m/(b_1 t) \tag{2.2}$$

where y denotes the consumer's gross income and p denotes the health insurance premium. b_1 and b_2 are co-payment rates and t is the producer price of medical services, so $b_1 t$ and $b_2 t$ stand for out-of-pocket prices. m refers to the maximum of co-payments that are collected using rate b_1.

The household problem is to maximize (2.1), subject to (2.2). This problem is complicated due to the endogeneity of the kink in the budget constraint, *i.e.* the household's choice for z defines the out-of-pocket price of his marginal unit of medical consumption. This makes it impossible to derive an analytical solution for the optimal co-payment rate and maximum without further simplification.

To make any progress, the next section simplifies the model by assuming that ε_z can take only two values: ε_{z_1} and ε_{z_2}. Effectively, this reduces the model to one with two types of consumers who face linear budget constraints with different slopes. Within this simplified setting, we are able to show that deductibles will never be part of the optimal co-payment scheme.

Subsequent sections elaborate the full model numerically, assuming that consumers are defined by their value of ε_z, which is drawn from a mixture of three distributions (two lognormal distributions and a mass point at zero). All consumers then face the same nonlinear budget constraint. Yet, the result that deductibles are sub-optimal is also found in this version of the model.

Before continuing, we make some additional comments about the interpretation of the model: about the resolution of uncertainty before patients make their consumption decision, about the role of risk sharing, about the impossibility for the health insurance industry to implement the first-best solution and about the absence of income heterogeneity.

As to the first point, our model assumes that health consumption decisions are made under certainty. This does certainly not hold true for all health demand decisions. The relevance of this assumption may be restricted however for two reasons. First, uncertainty about the future health status can affect the price of health care only for those who have not filled their co-payment maximum yet. Second, autocorrelation in medical shocks at the individual level reduces uncertainty for part of the patient population. The chronically ill are an obvious example. Feenberg and Skinner (1994) indicate that this type of autocorrelation is far from negligible.

The second point to make is that the model is about risk sharing, not distribution. *Ex post*, *i.e.* after the realization of health shocks, patients differ in their health status, their demand for health care, their consumption of non-medical services and their utility. The health insurance scheme then redistributes across patients. *Ex ante* however, patients are unknown whether they will be hit by a health shock and, if so, how intense this shock will be. *Ex ante*, therefore, all patients are equal. On an *ex ante* basis, thus, the insurance scheme implements risk sharing. As we assess the welfare implications of co-payment schemes on the basis of (*ex ante*) expected utility, it is risk sharing that, in addition to moral hazard, determines the welfare attached to different co-payment schemes.

We assume that only physicians can observe the health status of patients. They share this information with the patient but not with the insurer. Looking at the health care demand equation however, the health care demand exercised by patients reveals information about their health status. Still, this does not mean that the information on health status is public. There are a number of other factors that drive health care demand and that make identification of the health status of the patient impossible. We have chosen not to model these other factors explicitly as this would add little to our analysis.

Our model could easily be extended to one with heterogeneity across patients in terms of income. We have not done so in order to focus on moral hazard and risk, without having to bother about the issue of distribution. In an integral assessment of the effects of a typical co-payment scheme, such an extension would probably be very useful, however.

3 A simplified version

This section presents a simplified version of the model in the previous section. In particular, it specifies the frequency distribution for ε_z as a two-point mass distribution with mass points ε_{z_1} and ε_{z_2}, with $\varepsilon_{z_2} > \varepsilon_{z_1} > t$. As we will see, the last inequality ensures that in both states of nature patients consume a non-zero amount of medical care ($z_1, z_2 > 0$). The two values of ε_z correspond with the two parts of the budget constraint. In addition, the two values deviate sufficiently from the value of ε_z that corresponds with the kink in the budget constraint. To assess the precise implications of this must wait till the next section. In addition, this section puts ε_c at zero. In this case the utility function is linear in non-medical consumption and the income elasticity of health care demand equals zero.

$$u = c + \varepsilon_z z - \tfrac{1}{2}\varepsilon_m z^2 \tag{3.1}$$

Together, these assumptions imply that different co-payment rates apply to patients who are relatively healthy ($\varepsilon_z = \varepsilon_{z_1}$) and to patients who are relatively sick ($\varepsilon_z = \varepsilon_{z_2}$). In particular, the co-payment rate b_1 applies to patients with $\varepsilon_z = \varepsilon_{z_1}$ and b_2 applies to patients that have $\varepsilon_z = \varepsilon_{z_2}$. The budget constraints for the simplified version of our model thus read as follows:

$$\begin{aligned} c &= y - p - b_1 t z & \qquad \varepsilon_z &= \varepsilon_{z_1} \\ c &= y - p - m - b_2 t\left(z - \frac{m}{b_1 t}\right) & \qquad \varepsilon_z &= \varepsilon_{z_2} \end{aligned} \tag{3.2}$$

Optimization of (3.1) under the budget constraints in (3.2) gives us the equations for health care consumption:

$$\begin{aligned} z_1 &= \frac{-b_1 t + \varepsilon_{z_1}}{\varepsilon_m} \\ z_2 &= \frac{-b_2 t + \varepsilon_{z_2}}{\varepsilon_m} \end{aligned} \tag{3.3}$$

Inserting the demand for health care into the corresponding budget constraint (3.2) gives the demand for non-medical products. Substitution of the two types of demand into (3.1) gives us the indirect utility functions that expresses the maximum attainable utility as a function of the relevant co-payment rate:

$$v_1 = \frac{\varepsilon_m (y-p) + \frac{1}{2}(b_1 t)^2 + \frac{1}{2}\left(\varepsilon_{z_1}\right)^2 - \varepsilon_{z_1} b_1 t}{\varepsilon_m}$$

$$v_2 = \frac{\varepsilon_m (y - p - (1-b_2/b_1)m) + \frac{1}{2}(b_2 t)^2 + \frac{1}{2}\left(\varepsilon_{z_2}\right)^2 - \varepsilon_{z_2} b_2 t - \varepsilon_m (1-b_2/b_1)}{\varepsilon_m} \tag{3.4}$$

where v denotes indirect utility.

If we now normalize the size of the population to one and use E_1 and E_2 to denote the frequency of the respective population mass points, social welfare, to be denoted V, can be defined as follows:

$$V = E_1 v_1 + E_2 v_2 \tag{3.5}$$

Health insurance premiums are levied in order to finance the health care subsidies. Under the assumption of zero profits, the expression for health insurance premiums is straightforward:

$$p = E_1 (1-b_1) t z_1 + E_2 \left[(1-b_1) t \frac{m}{b_1 t} + (1-b_2) t \left(z_2 - \frac{m}{b_1 t} \right) \right] \tag{3.6}$$

The optimal co-payment structure now follows from the maximization of the social welfare function, subject to the constraint that the amount of health insurance premiums is exogenously given. Note that both V and p are functions of three policy instruments, namely b_1, b_2 and m. Without loss of generality, we can let m be a function of b_1 and b_2 by specifying m as $Rb_1/(b_1 - b_2)$, with R an arbitrary but positive constant. This makes V and p functions of two policy variables, b_1 and b_2. The point of maximum social welfare for a given amount of health insurance premiums now derives from $(\partial V/\partial b_1)/(\partial V/\partial b_2) = (\partial p/\partial b_1)/(\partial p/\partial b_2)$. We can elaborate this condition into the following expression:

$$\frac{(1-b_1)t}{\varepsilon_{z_1} - t} = \frac{(1-b_2)t}{\varepsilon_{z_2} - t} \quad \Rightarrow \quad \frac{(1-b_1)}{\varepsilon_{z_1} - t} = \frac{(1-b_2)}{\varepsilon_{z_2} - t} \tag{3.7}$$

Equation (3.7) states that the two co-payment rates are positively related. If we focus upon the case where the least healthy group is not subject to any co-payments, *i.e.* $b_2 = 0$, expression (3.7) demonstrates that the optimal co-payment rate for the most healthy group of consumers is strictly positive and smaller than one (recall that $\varepsilon_{z_2} > \varepsilon_{z_1} > t$). Hence, a co-payment scheme that features a deductible and zero co-payments for expenditure beyond the deductible cannot be optimal. Note that the size of the two groups is irrelevant, as is the amount of health insurance premiums (aggregate health insurance subsidies) for the form of the optimal co-payment structure.

Obviously, formula (3.7) could also have been obtained by directly applying the Ramsey rule for optimal commodity taxation. Sandmo (1987) shows that in case of two commodities that share the same producer price and that feature zero uncompensated cross-price elasticities, the following version of the Ramsey rule applies:

$$\tau_1\alpha_1 = \tau_2\alpha_2 \tag{3.8}$$

where τ_i i=1,2 is the subsidy rate on product i, *i.e.* the subsidy on product i in terms of the corresponding consumer price and α_i i=1,2 is the absolute value of the uncompensated own-price elasticity of demand for commodity i. Formally, $\tau_i = \left(t - t_{c_i}\right)/ t_{c_i}$ with t_{c_i} commodity i's consumer price and t the common producer price. The co-payment rate b_i i=1,2 is defined as t_{c_i} / t. Hence, equation (3.8) can also be formulated in terms of co-payment rates:

$$\frac{\left(1-b_1\right)}{b_1}\alpha_1 = \frac{\left(1-b_2\right)}{b_2}\alpha_2 \tag{3.9}$$

To elaborate this, use equation (3.3) to derive the following expressions for the (absolute value of the) price elasticity of the demand for medical services:

$$\begin{aligned} \alpha_1 &= \frac{b_1 t}{\varepsilon_{z_1} - b_1 t} \\ \alpha_2 &= \frac{b_2 t}{\varepsilon_{z_2} - b_2 t} \end{aligned} \tag{3.10}$$

Substitution of these expressions into the Ramsey rule (3.9) yields the following expression:

$$\frac{1-b_1}{\varepsilon_{z_1} - b_1 t} = \frac{1-b_2}{\varepsilon_{z_1} - b_1 t} \tag{3.11}$$

It can be easily derived that this is equivalent to expression (3.7).

4 A more realistic model

This section elaborates the full-fledged version of the model. We proceed in three steps (subsections 4.1, 4.2 and 4.3). We replace the two-spike distribution function of the parameter ε_z with a continuous distribution function. In contrast with the model in the previous section, the version in this section no longer imposes $\varepsilon_c = 0$ and puts $b_2 = 0$ from the start. The kinked nature of the budget constraint is preserved.

4.1 Optimization

We start to replace the two-spike distribution function of the parameter ε_z with a continuous distribution function. An important consequence of a continuous distribution function is that the level of medical spending that distinguishes the population groups now becomes endogenous. As we will see, the kink implies a bend (and discontinuity) in the health care demand equation. The condition that health care demand is nonnegative produces a second type of bend. The description of the solution procedure can therefore be split into three parts. The first two parts explain how the two bends are handled. The third part is necessary to rank the two bends. We start to describe the first part that describes the role of the first bend.

Part 1 optimization procedure

The kink in the budget constraint implies that for low levels of demand, the consumer pays a positive out-of-pocket price for his health care consumption, whereas he pays nothing for his consumption of services for high levels of demand.

$$\begin{aligned} c &= y - p - b_1 t z \qquad && 0 \le z \le m/(b_1 t) \\ c &= y - p - m \qquad && z \ge m/(b_1 t) \end{aligned} \tag{4.1}$$

The problem of the consumer is to maximize (2.1) under the constraint given by (4.1). The solution concept can be illustrated most clearly by considering it a three-step procedure (discussed also in Hausman (1985)).

The first two steps involve the solving of two hypothetical optimization problems, whereas the third step compares the utility levels of the two solutions. In the first step, the budget constraint in (4.1) is replaced by the following hypothetical downward-sloping linear constraint:

$$c = y - p - b_1 t z \tag{4.2}$$

Note that this constraint extrapolates the first part of budget constraint (2.2) for levels of z higher than $m/(b_1 t)$.

Optimization of (2.1) under this constraint produces the following expression for health care demand:

$$z_1 = \frac{-b_1 t(1-\varepsilon_c(y-p))}{\varepsilon_m + \varepsilon_c (b_1 t)^2} + \frac{\varepsilon_z}{\varepsilon_m + \varepsilon_c (b_1 t)^2} \tag{4.3}$$

where the subscript 1 indicates that (4.3) is the solution to the first hypothetical optimization problem. Inserting the demand for health care in (4.3) into the corresponding budget constraint (4.2) gives the demand for non-medical products. Substitution of the two types of demand into (2.1) gives the indirect utility level that corresponds to the first hypothetical optimization problem, v_1.

The second step is analogous to the first one, replacing the budget constraint by a horizontal line:

$$c = y - p - m \tag{4.4}$$

The corresponding expression for health care consumption reads as follows:

$$z_2 = \frac{\varepsilon_z}{\varepsilon_m} \tag{4.5}$$

Substitution of (4.5) and (4.4) into (2.1) gives the level of indirect utility that corresponds with the second hypothetical optimization problem, v_2.

The third step of the solution procedure involves the comparison of the indirect-utility levels that correspond to the two hypothetical optimization problems. The optimization problem that yields the highest level of utility is the solution to the consumer's problem.

As noted above, we consider the population to be heterogeneous in the need for health care. In particular, we assume the value of ε_z, which measures the utility of health care services relative to the utility of non-medical services, to take different values for different consumers. As Keeler *et al.* (1977) have shown, the non-convexity of the budget line implies that there is a particular value for ε_z, say ε_z^*, that has the feature that for all $\varepsilon_z > \varepsilon_z^*$, the consumer prefers the solution obtained in the second step above the solution obtained in the first step of the solution procedure and vice versa for $\varepsilon_z < \varepsilon_z^*$. For $\varepsilon_z = \varepsilon_z^*$, the consumer is indifferent between the two solutions (see Appendix A for the proof of this proposition and the elaboration of ε_z^*).

This setup has an interesting implication for the health demand curve, which we define here as the demand for medical consumption as a function of the health care parameter ε_z. This demand curve is increasing in a piecewise linear way. A discontinuity occurs at $\varepsilon_z = \varepsilon_z^*$. This reflects the discontinuous jump in the price and marginal utility of health care that occurs when the consumer reaches his co-payment maximum. Hence, there is a range of values for z which no patient chooses to consume. The discrete fall in the out-of-pocket price that occurs when ε_z passes the point ε_z^* implies a discrete increase in health care consumption. Also interesting is that the right-hand side of the demand curve is steeper than the left-hand part. Intuitively, the consumer is more responsive to a change in his need for medical services when he faces a zero out-of-pocket price.

Part 2 optimization procedure

As said, the health care demand curve has a second bend, which is due to the non-negativeness of health care demand. Unlike the co-payment maximum, the non-negativeness of health care demand does not produce a discontinuity in the demand function however. Rather, it specifies another critical value for ε_z, say ε_z^{**}, below which health care demand is zero. The calculation of ε_z^{**} proceeds exactly in the same way as that of ε_z^*. See Appendix A for further details.

We summarize the results obtained thus far by presenting the equations for health care demand and indirect utility.

$$
\begin{aligned}
z &= 0 & \varepsilon_z &\le \varepsilon_z^{**} \\
z &= \frac{-b_1 t(1-\varepsilon_c(y-p))}{\varepsilon_m + \varepsilon_c (b_1 t)^2} + \frac{\varepsilon_z}{\varepsilon_m + \varepsilon_c (b_1 t)^2} & \varepsilon_z^{**} &\le \varepsilon_z \le \varepsilon_z^* \\
z &= \frac{\varepsilon_z}{\varepsilon_m} & \varepsilon_z &\ge \varepsilon_z^*
\end{aligned}
\tag{4.6}
$$

$$
\begin{aligned}
v &= (y-p) - \tfrac{1}{2}\varepsilon_c (y-p)^2 & \varepsilon_z &\le \varepsilon_z^{**} \\
v &= \frac{\varepsilon_m (y-p) + \frac{1}{2}(b_1 t)^2 - \frac{1}{2}\varepsilon_c \varepsilon_m (y-p)^2 + \frac{1}{2}\varepsilon_z^2 - \varepsilon_z b_1 t(1-\varepsilon_c(y-p))}{\varepsilon_m + \varepsilon_c (b_1 t)^2} & \varepsilon_z^{**} &\le \varepsilon_z \le \varepsilon_z^* \\
v &= \frac{\varepsilon_m (y-p-m) - \frac{1}{2}\varepsilon_c \varepsilon_m (y-p-m)^2 + \frac{1}{2}\varepsilon_z^2}{\varepsilon_m} & \varepsilon_z &\ge \varepsilon_z^*
\end{aligned}
\tag{4.7}
$$

where ε_z^* and ε_z^{**} are defined in appendix A.

Part 3 optimization procedure

In general, $\varepsilon_z^* > \varepsilon_z^{**}$, *i.e.* small values of ε_z produce zero health care demand, medium values produce demand with co-payments below the maximum, and large values imply a level of health care demand the marginal unit of which is fully paid by the insurer. However, it can be shown that for sufficiently low values of the co-payment maximum *m*, the reverse case will occur: $\varepsilon_z^* < \varepsilon_z^{**}$. Then, one either does not consume medical services or consumes so much that the maximum amount of co-payments is to be paid. In this case, the critical values ε_z^* and ε_z^{**} are no longer relevant. Instead, a third critical value for ε_z, say ε_z^{***}, divides the demand function into two sections. The value for ε_z^{***} can be derived in a similar way as the critical values ε_z^* and ε_z^{**}. Hence, expressions (4.6) and (4.7) apply only if $\varepsilon_z^* > \varepsilon_z^{**}$. If $\varepsilon_z^* < \varepsilon_z^{**}$, demand and indirect utility are described by the following two equations:

$$\begin{aligned} z &= 0 & \varepsilon_z &\le \varepsilon_z^{***} \\ z &= \frac{\varepsilon_z}{\varepsilon_m} & \varepsilon_z &\ge \varepsilon_z^{***} \end{aligned} \tag{4.8}$$

$$\begin{aligned} v &= (y-p) - \tfrac{1}{2}\varepsilon_c (y-p)^2 & \varepsilon_z &\le \varepsilon_z^{***} \\ v &= \frac{\varepsilon_m (y-p-m) - \tfrac{1}{2}\varepsilon_c \varepsilon_m (y-p-m)^2 + \tfrac{1}{2}\varepsilon_z^2}{\varepsilon_m} & \varepsilon_z &\ge \varepsilon_z^{***} \end{aligned} \tag{4.9}$$

where ε_z^{***} is defined in appendix A.

Appendix A also derives the critical value for the co-payment maximum, denoted m^*, that specifies whether $\varepsilon_z^* < \varepsilon_z^{**}$ (and equations (4.6) and (4.7) apply), or whether $\varepsilon_z^* > \varepsilon_z^{**}$ (and equations (4.8) and (4.9) hold true).

4.2 Optimal co-payment policies

As indicated above, the continuous distribution for ε_z is a mixture of three distributions (two lognormal distributions and a mass point at zero). In particular, we explore a variant of the so called four-part model as applied by Duan *et al.* (1983) for the US and Van Vliet en Van der Burg (1996) for the Netherlands.

In its standard form, the four-part model describes the distribution of health care costs across individuals. In practice it appears that there are large numbers of people who do not consume health services in the period of analysis, that there are similarly large numbers of people who consume health services, but do not use any inpatient services, and that there are substantial differences between the health expenditure distributions of people who only incurred outpatient costs and those of people who have also been admitted to hospital. Furthermore, both latter distributions can reasonably well be described using a lognormal specification. The four-part approach combines these findings. First, it assumes that there is a non-zero probability that people will have zero health care costs. Secondly, it defines a non-zero probability that people who face positive costs will only consume outpatient services and a non-zero probability that their medical consumption includes inpatient services as well. Thirdly, it assumes that the health care spending of persons in groups E and I can be described by two separate lognormal distribution functions.

Our approach is based upon this four-part model, but relates it to medical need rather than to health care expenditure. It assumes that there is a given non-zero probability π_0 of having zero medical need. There is a also given non-zero probability π_E that a positive need is a draw from a distribution that describes need without inpatient services. Finally, there is a known non-zero probability $\pi_1 = 1 - \pi_E$ that the need for health care is a draw from the distribution of need that also includes inpatient services. Obviously, it holds that $\pi_0 + (1-\pi_0)\pi_E + (1-\pi_0)\pi_I = 1$. The non-zero medical needs of patients for medical services excluding inpatient services (denoted as group E) and for medical services including inpatient services (denoted as group I) are described by two separate lognormal distribution functions.

The demand for health care within the two groups obeys equations (4.6) and (4.8) in the previous section. It will be clear that when need is zero, demand for health care services is also zero. The opposite is generally not true, however. As we will see, patients may decide not to consume medical services in case of a positive need (this happens for those who have $\varepsilon_z \le \varepsilon_z^{**}$ (in case equation (4.6) applies) or $\varepsilon_z \le \varepsilon_z^{***}$ (in case equation (4.8) applies)). This implies that the probability of incurring zero costs is endogenous and may exceed the exogenous probability of zero need, π_0.

Summing up, our version of the four-part model implies that there is a non-zero probability $(1-\pi_0)\pi_E$ that $\ln(\varepsilon_z) \sim N(\mu_E, \sigma_E^2)$, $\varepsilon_m = \varepsilon_{m,E}$ and $t = t_E$, a non-zero probability $(1-\pi_0)\pi_I$ that $\ln(\varepsilon_z) \sim N(\mu_I, \sigma_I^2)$, $\varepsilon_m = \varepsilon_{m,I}$ and $t = t_I$, and a nonzero probability π_0 that ε_z equals zero (the values of ε_m and t are irrelevant in case $\varepsilon_z = 0$).

This setup implies the following expression for aggregate indirect utility, which serves as our measure of social welfare:

$$V = \pi_0\left((y-p) - \tfrac{1}{2}\varepsilon_c (y-p)^2\right) + (1-\pi_0)\pi_E V_E + (1-\pi_0)\pi_I V_I \qquad (4.10)$$

with:

$$V_i = \left[G\left(\varepsilon_z^{**}\right)\left((y-p) - \tfrac{1}{2}\varepsilon_c (y-p)^2\right) + \left(G\left(\varepsilon_z^{*}\right) - G\left(\varepsilon_z^{**}\right)\right) E\left(v \mid \varepsilon_z^{**} \le \varepsilon_z \le \varepsilon_z^{*}\right)\right]_i$$

$$+ \left[\left(1 - G\left(\varepsilon_z^{*}\right)\right) E\left(v \mid \varepsilon_z \ge \varepsilon_z^{*}\right)\right]_i$$

$$m \ge m_i^* \tag{4.11}$$

$$= \left[G\left(\varepsilon_z^{***}\right)\left((y-p) - \tfrac{1}{2}\varepsilon_c (y-p)^2\right) + \left(1 - G\left(\varepsilon_z^{***}\right)\right) E\left(v \mid \varepsilon_z \ge \varepsilon_z^{***}\right)\right]_i$$

$$m \le m_i^*$$

$$i = E, I$$

Here, *G(.)* denotes the distribution function of ε_z. The three conditional expectation variables $E(v|.)$ in equations (4.10) and (4.11) are derived from equations (4.7) and (4.9).

The structure of the expression for health insurance premiums resembles that of the expression for indirect utility:

$$p = (1-\pi_0)\pi_E p_E + (1-\pi_0)\pi_I p_I \tag{4.12}$$

with:

$$p_i = \left(G\left(\varepsilon_z^{*}\right) - G\left(\varepsilon_z^{**}\right)\right)(1-b_1) t E\left(z \mid \varepsilon_z^{**} \le \varepsilon_z \le \varepsilon_z^{*}\right) + \left(1 - G\left(\varepsilon_z^{*}\right)\right)\left[t\, E\left(z \mid \varepsilon_z \ge \varepsilon_z^{*}\right) - m\right]$$

$$m \ge m_i^*$$

$$= \left(1 - G\left(\varepsilon_z^{***}\right)\right)\left[t\, E\left(\varepsilon_z \ge \varepsilon_z^{***}\right) - m\right] \tag{4.13}$$

$$m \le m_i^*$$

$$i = E, I$$

The three conditional expectation variables $E(z \mid .)$ in equations (4.12) and (4.13) can be derived from the expressions in equations (4.6) and (4.8).

After substitution of the equation for health insurance premiums into the indirect utility function, *V*, we are left with an expression for social welfare that is a function of two policy instruments, b_1 and *m*. The function is very complicated, as all critical values ε_z^*, ε_z^{**}, ε_z^{***} and m^* are endogenous, and can not be solved analytically. Hence, we resort to numerical simulations. In particular, we make calculations for a large number of co-payment schemes that differ along two dimensions: the co-payment rate and the co-payment maximum. The co-payment scheme that corresponds with the highest level of social welfare can then be said to be optimal.

5 Calibration of model parameters

Before the year 2006, the Dutch public health insurance scheme was accessible only for those with income below a certain threshold; high incomes had to buy insurance on the private market. In 2006, the public insurance scheme was extended to the whole population and the distinction that had long characterized the Dutch market for health insurance came to an end.

We calibrate our model on data for 2002, the year for which the most recent data are available. In that year, the public insurance scheme did not feature any co-payments. Absent co-payments, it is impossible to calibrate the model. Hence, we calibrate the model on the population of privately insured, who traditionally have faced co-payments, particularly in the form of deductibles. At the end of the next section, we will say something about the implications of our model for the population of the former publicly insured.

Our calibration procedure starts with a specification of the distribution function for ε_z. Van Vliet and Van der Burg (1996) have analyzed the properties of the distribution of health care expenditure of people with and without consumption of inpatient services, based on Dutch cross section data for 1991–1994. We use their estimates of the coefficients of variation of the lognormal distribution functions to obtain corresponding values for ε_z. From these values, it is easy to derive estimates of the standard deviations σ_E and σ_I of the normal parts themselves. The probability of zero need for health care services, π_0, has been calibrated such that the calculation in the model of the probability of zero health expenditure matches the data. The value of the probability to consume inpatient health services, π_I, has been approximated by the number of hospital admissions per insured patient with positive health care expenditure.[7]

This leaves us with five parameters that remain to be identified: ε_c, $\varepsilon_{m,E}$, $\varepsilon_{m,I}$, μ_E and μ_I. These are calibrated simultaneously by using information on the coefficient of relative risk aversion (CRRA) for non-medical products, the level of average demand of those with and those without inpatient services and the insurance effect[8] corresponding to the demand of these two groups. The services covered are pharmaceuticals and services delivered by general practitioners, dentists, physiotherapists and hospitals (inpatient and outpatient services). Services in groups E and I are different, so prices do not coincide either. We use the estimates of Van Vliet (1998) to compute the insurance effects for the groups with positive health care expenditure, both without hospital admissions (1.23, group E) and with both outpatient and inpatient expenditure (1.03, group I). Note that these estimates are in line with estimates in the RAND Health Insurance Experiment (Newhouse *et al.* (1993)). For the CRRA for non-medical products, we choose a value of 5.

[7] This assumes that patients who used inpatient hospital services, had only one admission per year. This is probably not too far off from reality.

[8] The insurance effect is defined as the ratio of the demand of a fully insured patient (with a zero co-payment maximum) and an uninsured patient.

Before we proceed, a caveat is in order. Whereas the empirical literature finds the four-part model to provide a fairly well description of health expenditures, we assume the four-part model characterizes ε_z , the parameter that we interpret as a measure of the need of health care. Were the two variables proportional to each other, the difference in interpretation would be trivial. However, this is not the case for two reasons.

First, log-normality of the distribution functions for ε_z means that the distribution function for z is not completely lognormal, according to our model. Indeed, our model implies that no patient will choose to consume medical services at an expenditure level that is sufficiently close to the co-payment maximum. This means that the cumulative distribution function of medical expenditure features a flat trajectory around the co-payment maximum. Second, the proportionality factor that translates ε_z into z , $\partial z / \partial \varepsilon_z$, is different for the parts of the distribution function for z below and above the co-payment maximum.

In our application, these two issues appear not to be very relevant, though. The region of values of z that have frequency zero range from 2.20 to 4.42 in the case of group E for example, whereas average health care consumption in this group is much larger, namely 14.47. Next, the two proportionality factors referred to above differ less than one percent (0.40 and 0.398 respectively). To explore this further, we will perform a sensitivity analysis in order to see the impact of alternative assumptions on the parameter values of the distribution functions for ε_z .

We do not calibrate our model on estimates of the price elasticity of health care demand. In our model, the price elasticity of demand is endogenous, depending among others on the value of ε_z (which measures the need for health care), the co-payment rate b and the maximum m. The unconditional average price elasticity for those who consume outpatient services only (E services) is -0.39. The unconditional average price elasticity for those who consume both inpatient and outpatient services (I group) equals zero as in this case equation (4.8) applies (see section 4.1, Part 3 optimization procedure). The average of these two price elasticities equals -0.36, which is somewhat higher than the average estimate -0.20 of Newhouse *et al.* (1993). We did not calibrate on the income elasticity of health care demand either. This income elasticity is related to the CRRA for non-medical products. Having chosen a value for the latter, the income elasticity of health care demand follows endogenously from the model. The average income elasticity of health care demand is 1.35 for E-type services and 0.0 for I-type services. The reason for the zero income elasticity is the same as the reason for the zero price elasticity: a negligible part of the patients that consume inpatient services has expenditure that qualifies for less than the maximum of co-payments. On average, we calculate an income elasticity of 1.24. This value for the aggregate income elasticity is close to estimates that analyze health care demand on a macroeconomic level (Gerdtham *et al.* (1992) find an aggregate income elasticity of 1.27).[9] Finally, our value for the CRRA for health care that can be derived from the calibration is 4.5 which deviates little from the value of 5 used by Ellis (1986).[10]

Table 5.1 Validation of model parameters: data

	Privately insured			Publicly insured		
	Group E	Group I	Total	Group E	Group I	Total
Probability of zero expenditure (%)			23.1			17.9
Probability of positive expenditure on outpatient services only (%)			70.7			74.7
Probability of positive expenditure on outpatient and inpatient services (%)			6.2			7.4
Insurance effect	1.23	1.03		1.23	1.03	
CRRA, non-health products			5.0			5.0
Income per patient (euro)			39,235			15,093
Average demand health care services	14.5	24.9		21.3	29.7	
Producer price health care services (euro)	24.4	239.4		17.0	230.7	
Coefficient of variation health care costs	2.03	1.11		1.71	1.20	
Co-payment maximum (euro)			64.2			0.0
Co-payment rate (%)			100			0.0

Tables 5.1 and 5.2 summarize the calibration of our model. Parameter estimates for the publicly insured are also included, to be discussed below. A salient aspect of the calibration is the skewed nature of need and demand. If a person spends on health care, there is a chance of about 1 to 10 that this spending includes inpatient hospital services. However, the average volume of health care consumption of inpatient services is about double that of outpatient services and the price of inpatient services is about ten times as large as the price of outpatient services. Combined, spending on inpatient services is larger than that on outpatient services.

Also interesting is the group of persons that do not consume any health care. About a quarter of the population of privately insured has zero expenditure. Of this, only 2.5% has zero medical need. For the most part, people that choose to have zero spending are people with positive need

[9] This value for the aggregate income elasticity of demand is much higher than usually found in microeconometric studies. However, we argue that estimates of the income elasticity found in macroeconometric studies are relevant here. Income elasticity estimates found in microeconometric studies apply to changes in income for an individual patient. The income elasticities found in macroeconometric studies apply to changes in income for the whole population of patients. The co-payment schemes in our study differ in the amount of insurance premiums and this in the disposable income of all patients. In addition, we should point out that the income elasticity of health care demand and the coefficient of relative risk aversion with respect to non-medical produts are intimately related. An income elasticity of the order of magnitude found in microeconometric studies would imply a value for the CRRA well below unity (0.09), which must be considered highly unrealistic.

[10] A number of studies (*e.g.* Manning and Marquis (1996)) report estimates of the coefficient of absolute risk aversion. These data are difficult to use in our study, as their value is not unit-independent.

for outpatient services that is so small that the benefits from medical intervention are less than the costs involved.

Table 5.2 Validation of model parameters: results

	Privately insured			Publicly insured		
	Group E	Group I	Total	Group E	Group I	Total
Probability of zero need (%)			2.5			8.6
Probability of positive need for E-type services (%)			89.5			82.4
Probability of positive need for I-type services (%)			8.0			9.0
ε_c			$2.42\ 10^{-5}$			$7.16\ 10^{-5}$
ε_m	2.5	140.2		0.8	200.4	
μ	2.3	7.6		2.8	8.2	
σ	1.28	0.90		1.17	0.94	
Average need per patient ($E(\varepsilon_z)/\varepsilon_m$)	15.3	24.9		23.6	29.7	

To identify model parameters for formerly publicly insured patients, we assume that insurance effects of privately insured patients also apply to this group. Also, the value of CRRA for non-health commodities and services is assumed to be the same for the two groups of insured. As Table 5.1 shows, income is considerably smaller, but average health care demand is larger. This is due to a larger need for medical services (Table 5.2) and a zero price (Table 5.1).

6 Numerical calculations

The approach we adopt to find out which co-payment scheme is optimal in the full-fledged model is to search numerically for the combination of policy instruments (b, m) that maximizes social welfare, as defined in equation (4.10). If the value of b that corresponds with optimal co-payment policies equals one, a deductible scheme can be said to be optimal.

Tables 6.1 to 6.5 present calculations for a number of combinations of b and m. In particular, they display effects for 11 different values for b (running from 0% to 100% in steps of 10%) and 12 different values for m (running from euro 0 to euro 2000 in steps of euro 200 plus infinity). Table 6.1 displays the effects upon $E(z)$, the average level of medical consumption. The cells with a zero co-payment rate and a non-zero co-payment maximum are left blank; the same applies to the cells with positive co-payment rate and zero co-payment maximum. In both cases, there is full insurance and the figure for (b = 0, m = 0) applies.

Table 6.1 Expected volume of health care consumption of privately insured patients

Rate (%)	Co-payment maximum (euro)											
	0	200	400	600	800	1000	1200	1400	1600	1800	2000	∞
0	14.9											
10		14.8	14.8	14.8	14.7	14.7	14.7	14.7	14.7	14.7	14.7	
20		14.6	14.6	14.6	14.6	14.6	14.6	14.6	14.6	14.6	14.6	
30		14.5	14.4	14.4	14.4	14.4	14.4	14.4	14.4	14.4	14.4	
40		14.4	14.3	14.2	14.2	14.2	14.2	14.2	14.2	14.2	14.2	
50		14.3	14.1	14.1	14.0	14.0	14.0	14.0	14.0	14.0	14.0	
60		14.2	14.0	13.9	13.9	13.9	13.9	13.9	13.8	13.8	13.8	
70		14.1	13.9	13.8	13.8	13.7	13.7	13.7	13.7	13.7	13.7	
80		14.1	13.8	13.7	13.6	13.6	13.6	13.6	13.5	13.5	13.5	
90		14.0	13.7	13.6	13.5	13.5	13.4	13.4	13.4	13.4	13.4	
100		14.0	13.7	13.5	13.4	13.3	13.3	13.3	13.3	13.2	13.2	13.2

As might be expected, $E(z)$ is declining in both the co-payment rate and the co-payment maximum. The impact of variations in the co-payment rate seems strongest. Looking at the corner columns and rows, the parameters interact strongly, however. Hence, it is difficult to calculate the exact contribution of the two parameters. Moving from full insurance ($b = m = 0$) to zero insurance (b = 100, $m = \infty$[11]) reduces average medical consumption by about 11.5 percent. This effect is not trivial, although smaller than expected. It is clearly far below the

[11] Approximated by 10^8 euro.

almost 50% reduction of expenditures that can be calculated on the basis of the Health Insurance Expreiment (Zweifel and Manning (2000)).

Table 6.2 gives the effects upon $E(c)$, the average level of non-medical consumption. Here also, the qualitative effects are intuitive. The higher the rate or the maximum of co-payments, the higher is average non-medical consumption. The reason is that any reduction in average health care consumption allows for a cut in health insurance premiums.

On the whole, the effects are pretty small. Going from full insurance to no insurance would raise average non-medical consumption with a meagre 0.2%. The reason is clear: the effects on average medical consumption are small and the spending on non-medical consumption dwarfs that on medical consumption.

Table 6.2 Expected volume of non-medical consumption of privately insured patients

Rate (%)	Co-payment maximum (euro)											
	0	200	400	600	800	1000	1200	1400	1600	1800	2000	∞
0	32,076											
10		32,081	32,081	32,081	32,081	32,081	32,082	32,082	32,082	32,082	32,082	
20		32,085	32,086	32,086	32,086	32,086	32,086	32,087	32,087	32,087	32,087	
30		32,088	32,090	32,090	32,091	32,091	32,091	32,091	32,091	32,092	32,092	
40		32,091	32,093	32,094	32,095	32,095	32,095	32,096	32,096	32,096	32,096	
50		32,093	32,096	32,098	32,098	32,099	32,099	32,100	32,100	32,100	32,101	
60		32,095	32,099	32,101	32,102	32,103	32,103	32,104	32,104	32,104	32,105	
70		32,096	32,102	32,104	32,105	32,106	32,107	32,107	32,108	32,108	32,109	
80		32,097	32,104	32,107	32,108	32,110	32,110	32,111	32,111	32,112	32,112	
90		32,098	32,106	32,109	32,111	32,113	32,114	32,114	32,115	32,115	32,116	
100		32,099	32,107	32,112	32,114	32,115	32,117	32,117	32,118	32,119	32,119	32,129

The impact of insurance can be highlighted by computing the implicit transfers from patients in the E group (patients that consume only outpatient services) to those in the I group (patients that consume also inpatient services). These implicit transfers are defined as the premium actually paid minus the hypothetical amount that patients in the E group would pay if the I group would be excluded from health insurance. The implicit transfers are at maximum in case of full insurance. The amount of 536 euro per E patient corresponds to 6,168 euro per patient in the I group (89.4% of the population falls into the E group, 9% into the I group, 2.6% has zero need). The latter amount is about 15% of the average patient income, 39,235 euro, which is sizeable.

In general, implicit transfers are decreasing, both in terms of the co-payment rate and in terms of the co-payment maximum. Obviously, downsizing insurance decreases implicit transfers. For small values of the co-payment rate, implicit transfers are an increasing function of the co-payment rate, however. This has to do with the differential price elasticity of the demand for inpatient and outpatient care. An increase of the co-payment rate does not have a big effect on medical consumption of I-type services, whereas patients in the I group share in the fall in insurance premiums that is due to lower medical consumption by patients in the E group. Hence, implicit transfers from the E group to the I group increase upon an increase in co-payment rates.

In the limit case (b=100%, m= ∞) implicit transfers are zero. Interestingly, with a co-payment maximum as high as 2000 euro, the implicit transfer per patient of E-type services is still 414 euro. This again reflects that the consumption of I-type services is very price-inelastic.

Table 6.3 Average implicit income transfers from E to I patients per patient in group E

Rate (%)	Co-payment maximum (euro)											
	0	200	400	600	800	1000	1200	1400	1600	1800	2000	∞
0	536											
10		524	513	506	500	496	493	491	490	488	487	
20		526	512	500	490	482	475	469	464	460	456	
30		528	514	500	488	477	467	458	451	444	438	
40		529	515	501	488	476	464	453	444	435	427	
50		530	517	503	489	476	463	451	440	430	421	
60		530	518	504	490	477	464	451	439	428	417	
70		531	519	506	492	478	464	451	439	426	415	
80		531	520	507	493	479	465	452	439	426	414	
90		531	521	508	494	480	466	453	439	426	414	
100		532	521	509	495	482	468	454	440	427	414	0

Table 6.4 Social welfare of privately insured patients

Rate(%)	Co-payment maximum (euro)											
	0	200	400	600	800	1000	1200	1400	1600	1800	2000	∞
0	965.9											
10		975.5	975.3	974.8	974.3	973.8	973.3	972.9	972.5	972.1	971.8	
20		982.6	982.8	981.9	980.7	979.4	978.0	976.7	975.5	974.2	973.1	
30		987.6	988.8	987.8	986.1	984.1	982.0	979.8	977.6	975.4	973.3	
40		990.9	993.4	992.5	990.6	988.0	985.2	982.2	979.2	976.0	972.9	
50		992.8	996.5	996.0	994.0	991.1	987.7	984.0	980.1	976.2	972.1	
60		993.6	998.5	998.4	996.3	993.2	989.3	985.1	980.5	975.8	970.9	
70		993.7	999.4	999.7	997.6	994.3	990.2	985.4	980.3	974.8	969.2	
80		993.1	999.4	**1000.0**	998.0	994.6	990.1	985.0	979.3	973.3	966.9	
90		992.1	998.7	999.5	997.6	994.0	989.4	983.9	977.7	971.1	964.2	
100		990.8	997.4	998.3	996.4	992.7	987.8	982.0	975.5	968.4	960.9	0.0

Table 6.4 displays social welfare. For reasons of presentation only, values in the table are linear transformations of computed welfare. It turns out that the (*b*, *m*) combination of (80%, 600 euro) yields the highest level of social welfare (displayed boldly in Table 6.4). That the optimal co-payment rate and co-payment maximum have intermediate values, reflects the trade-off between the welfare loss from the moral hazard distortion and the welfare gain from risk reduction that health insurance brings about. (b, m) combinations to the upper left of the optimum (denoted as $(\hat{b}, \hat{m})$) imply lower social welfare as they feature too much moral hazard. Similarly, (b, m) combinations to the lower right of $(\hat{b}, \hat{m})$ imply lower social welfare as they correspond with too little risk sharing.

As a tool to produce an accurate estimate of the optimal co-payment scheme, the search procedure that produced table 6.4 is too global: we have to resort to a finer grid search to find the globally optimal co-payment scheme. Therefore, we declined the step sizes of b and m to 2% and 100 euro respectively and arrived at (b = 78%, m =530) as the optimal co-payment scheme. The corresponding level of social welfare equals 1,000.2. A further reduction of the step size of b and m produced negligible changes (differences in expected utility less than 10^{-3} %). A further refinement of the search procedure is therefore not necessary.

Full insurance dominates no insurance. This result differs strongly with that in Manning and Marquis (1996). That analysis found the case without any insurance to dominate the case of full insurance. The income elasticities used in the two studies may offer an explanation. Manning and Marquis (1996) use an income elasticity of 0.22, which is considerable smaller than ours and would imply a much lower CRRA for non-medical products than we have used. In addition, that analysis assumes the price elasticity of medical consumption to be a constant, whereas ours decreases with a deterioration of the health status of the patient. Moreover, in general one may expect large differences for extreme cases like full insurance and no insurance, since calculation of the welfare levels of these cases relies on an extrapolation of parameter values far beyond the range of data that were used to estimate them.

For a meaningful comparison of welfare effects, we resort to the corresponding compensating variations. These are defined as follows. Formally, if we express social welfare as a function of the rate and maximum of co-payments and the level of patient income, say $V(b, m, y)$, we calculate the compensating variation for full insurance, $\tilde{y}_{FI}$, from the condition $V(0{,}0, y + \tilde{y}_{FI}) = V(\hat{b}, \hat{m}, y)$. Similarly, we calculate the compensating variation for zero insurance, $\tilde{y}_{ZI}$, from $V(100\%{,}40000, y + \tilde{y}_{ZI}) = V(\hat{b}, \hat{m}, y)$.[12] The generalization to the definition of compensating variation as a function of the rate and maximum that define a typical co-payment scheme is obvious then.

[12] If *m* exceeds 40000 euro per person, the condition $\varepsilon_c(y + \tilde{y}_{ZI} - p) < 1$ is violated (see also footnote 5). Hence, we cannot calculate a compensating variation for the case without any insurance. The reason is that utility is quadratic in income. Hence, there is an upper limit to the utility gain that a compensating variation of income can achieve,. If the differential utility with the optimum ($(\hat{b}, \hat{m})$) is too large (as it is for the case of zero insurance), there is no value for the compensating variation that would bridge the utility differential.

Table 6.5 repeats Table 6.4, but now displaying compensating variations rather than social welfare levels. Full insurance is seen to imply a utility loss of only 18 euro per person, equivalent to an income drop of less than 0.1 percent. The compensating amount in case of zero insurance amounts to 419 euro per patient. This corresponds to 1.1% of average income per patient. This is extremely little.

As noted earlier, until January 1 2006 there were two insurance schemes in the Netherlands, a set of private insurance schemes and a public insurance scheme. Using the values in Table 5.1 and 5.2, it is possible to evaluate the optimal insurance scheme for publicly insured patients. The optimum turns out to be (68%, 320 euro), which is rather close to the outcome for the privately insured (78%, 530 euro). On average, we arrive at an optimum of (72%, 380 euro) for all insured patients.

Table 6.5 Compensating variations for different (b,m) combinations, privately insured (euro)

Rate (%)	Co-payment maximum (euro)											
	0	200	400	600	800	1000	1200	1400	1600	1800	2000	40000
0	18.23											
10		13.13	13.23	13.48	13.76	14.04	14.29	14.52	14.73	14.91	15.07	
20		9.32	9.22	9.70	10.35	11.05	11.76	12.47	13.14	13.79	14.41	
30		6.63	5.99	6.53	7.43	8.51	9.65	10.83	12.01	13.17	14.32	
40		4.87	3.55	4.00	5.05	6.41	7.92	9.52	11.17	12.84	14.51	
50		3.85	1.86	2.12	3.23	4.79	6.60	8.57	10.64	12.77	14.93	
60		3.40	0.82	0.86	1.98	3.67	5.72	8.00	10.44	12.98	15.60	
70		3.39	0.34	0.18	1.26	3.05	5.28	7.82	10.58	13.50	16.53	
80		3.70	0.33	0.00	1.05	2.90	5.28	8.04	11.08	14.33	17.73	
90		4.23	0.72	0.27	1.29	3.20	5.71	8.65	11.93	15.47	19.21	
100		4.93	1.42	0.92	1.94	3.90	6.53	9.64	13.14	16.94	20.99	418.76

7 Sensitivity analysis

How robust are our results with respect to perturbations in important parameters? Three of them relate directly to the moral hazard and risk reduction components of expected utility. These are ε_z , ε_m and the CRRA. The last parameter is directly linked to the parameter ε_c . We vary the values of ε_m , the CRRA and the variance of ln(ε_z) between 25% and 175% of their benchmark values, in steps of 25%. Changes in the variance of ln(ε_z) are equivalent to changes in the coefficient of variation of the distribution of ε_z itself. We adopt step sizes in the co-payment rate of 10% and 200 euro in the co-payment maximum, like in Table 6.4. Tables 7.1, 7.2 and 7.3 contain the results: the optimal co-payment rate and the co-payment maximum. The last row displays compensating variations relating to the optimum in each column and the corresponding situation of full insurance. Note that in Tables 7.1 and 7.3 we vary the parameters for both the E and I group simultaneously. Finally, in Table 7.4 we report the results for changes in average patient income. Here we have chosen a different range in order not to violate the restriction on income $\varepsilon_c\ (y\text{-}p) < 1$ (see also note 5).

Table 7.1 Optimal (b,m) combinations and compensating income variations for different values of the standard deviation of ln(ε_z)

	Standard deviation of ln(ε_z)						
E group	0.32	0.64	0.96	1.28 (base)	1.60	1.92	2.24
I group	0.23	0.45	0.68	0.90 (base)	1.12	1.34	1.57
Co-payment rate (%)	90%	80%	80%	80%	70%	70%	60%
Maximum (euro)	800	800	600	600	400	200	200
Compensating variation	24.56	22.04	20.00	18.23	15.86	12.67	9.94

Table 7.2 Optimal (b,m) combinations and compensating income variations for different values of the CRRA

	CRRA						
	1.25	2.50	3.75	5.00	6.25	7.50	8.75
Co-payment rate (%)	80%	80%	80%	80%	80%	80%	80%
Maximum (euro)	800	600	600	600	400	400	400
Compensating variation	39.96	28.25	22.11	18.23	15.56	13.86	12.55

Table 7.3 **Optimal (b,m) combinations and compensating income variations for different values of ε_m, both E and I group**

	ε_m						
E group	0.62	1.25	1.87	2.49 (base)	3.11	3.74	4.36
I group	35.04	70.08	105.12	140.16 (base)	175.20	210.24	245.28
Co-payment rate (%)	70%	80%	80%	80%	80%	80%	80%
Maximum (euro)	600	600	600	600	600	400	400
Compensating variation	62.08	35.41	24.31	18.23	14.39	11.88	10.11

Table 7.4 **Optimal (b,m) combinations and compensating income variations for different values of average patient income**

	Average income per patient (euro)						
	19618	23541	27465	31388	35312	39235(base)	43159
Co-payment rate (%)	80%	80%	80%	80%	80%	80%	80%
Maximum (euro)	800	800	600	600	600	600	400
Compensating variation	48.80	43.52	37.85	31.87	25.38	18.23	10.82

Given the wide range in parameter values, the optimal co-payment rate is fairly stable. Its value varies between 70% and 90%. This implies that our conclusion that $b < 100\%$ is quite robust. The optimal value of m is more sensitive, in particular with respect to downward changes in model parameters. When differences in health status between individuals decline, risk aversion is less important and the optimal co-payment maximum and compensating income amount increase. The same happens when the CRRA is reduced, as was to be expected. Table 7.2 demonstrates that the impact of the CRRA parameter upon the optimal co-payment maximum is fairly large: the optimal co-payment maximum increases from 400 to 800 euro when CRRA drops from a value of 5 to a value of 1.25. Simulations of still lower values of the CRRA (not shown here) indicate even higher values for the co-payment maximum. The high sensitivity of the optimal co-payment maximum to the value of the coefficient of risk aversion was also noted by Manning and Marquis (1996). Finally, a fall in the value of ε_m yields a lower price elasticity of health care demand and thus increases moral hazard. This also leads to higher values for the optimal co-payment maximum. The impact of this parameter is far less large than that of the parameter that measures risk aversion. Table 7.4 indicates that changes in average income have a small effect on the optimal co-payment rate and a large effect on the optimal co-payment maximum.

8 Concluding remarks

The most important results from our analysis are the following: i) within the class of co-payment systems studied, a deductible scheme is suboptimal ii) the optimal co-payment scheme features a co-payment rate of about 80% and iii) the optimal co-payment scheme features a co-payment maximum of about 600 euro. The robustness of these three results is quite different.

The sub-optimality of deductible schemes is quite robust. As shown, it holds true both in a stylized world in which it derives from the Ramsey rule of optimal commodity taxation and in a world that includes much more realism. However, the other two results depend on the numerical setting that we adopted and are therefore less general. Although the optimal co-payment rate is roughly stable at about 80%, the sensitivity analysis that we performed made clear that the results regarding to co-payment maximum rely rather heavily on the degree of risk aversion, which is a parameter of which estimates have large confidence intervals.

Three factors warn against immediate application of our results by policymakers. First, our calculations did not include administrative costs, which can be different for deductible schemes than for alternative co-payment systems. Second, preferences for policymakers need not coincide with those of households. In particular, we think the risk aversion of policymakers may be higher than that of households. Thirdly, a caveat is that we did not explore the consequences of other types of utility functions. This underlines that our results cannot be more than indicative of the answers that a more comprehensive analysis would produce.

Appendix A Critical values for the health care need parameter

The proof of the proposition that there is a particular value for ε_z, ε_z^* that has the feature that i) for $\varepsilon_z > \varepsilon_z^*$ ($\varepsilon_z < \varepsilon_z^*$), the consumer prefers the solution obtained in the second (first) step above the solution obtained in the first (second) step of the solution procedure and ii) for $\varepsilon_z = \varepsilon_z^*$, the consumer is indifferent between the two solutions, is as follows. First, note that v_1 and v_2, the indirect utilities corresponding to the two hypothetical optimization problems, can be viewed as functions of ε_z. To stress this relationship, denote these functions as $v_1(\varepsilon_z)$ and $v_2(\varepsilon_z)$. Next, define the difference between the two functions, $v_{1-2}(\varepsilon_z)$, as $v_1(\varepsilon_z) - v_2(\varepsilon_z)$. Given the specification of the utility function in (2.1) and the optimal solutions in (4.7) and (4.9), the function $v_{1-2}(\varepsilon_z)$ is quadratic in ε_z. Solving $v_{1-2}(\varepsilon_z) = 0$ yields two solutions for ε_z, of which one is positive. Call this solution ε_z^*. It can be derived that $v_{1-2}(\varepsilon_z)$ is decreasing in ε_z for all $\varepsilon_z > 0$. Hence. $v_{1-2}(\varepsilon_z) > 0$ for $0 < \varepsilon_z < \varepsilon_z^*$ and $v_{1-2}(\varepsilon_z) < 0$ for $\varepsilon_z > \varepsilon_z^*$.

The general expression for the equation $v_{1-2}(\varepsilon_z) = 0$ reads as follows:

$$\nu_a \varepsilon_z^2 + \nu_b \varepsilon_z + \nu_c = 0 \tag{A.1}$$

with

$$\begin{aligned} \nu_a &= (b_1 t)^2 \left(\frac{\varepsilon_c}{2\varepsilon_m} \right) \\ \nu_b &= b_1 t (1 - \varepsilon_c (y - p)) \\ \nu_c &= -\varepsilon_m m (1 - \varepsilon_c (y - p)) - \tfrac{1}{2} \varepsilon_m \varepsilon_c m^2 - \tfrac{1}{2} (b_1 t)^2 (1 - \varepsilon_c (y - p - m))^2 \end{aligned} \tag{A.2}$$

The positive root of (A.1) obeys:

$$\varepsilon_z^* = \frac{-\nu_b}{2\nu_a} + \frac{1}{2\nu_a} \sqrt{\nu_b^2 - 4\nu_a \nu_c} \tag{A.3}$$

As the maximum of expression (A.1) is achieved for $\varepsilon_z = \dfrac{-\nu_b}{2\nu_a}$, equation (A.1) is decreasing in ε_z whenever $\varepsilon_z > \dfrac{-\nu_b}{2\nu_a}$.

The procedure we adopt to derive expressions for ε_z^{**} and ε_z^{***} is similar to the one we use to derive ε_z^*. The derivation of the expression for ε_z^* uses the indirect utility functions corresponding to the two hypothetical optimization problems mentioned in the text (v_1 and v_2, or v for $\varepsilon_z \le \varepsilon_z^{**}$ and v for $\varepsilon_z^{**} \le \varepsilon_z \le \varepsilon_z^*$ in equation (4.7) respectively). The derivation of

ε_z^{**} combines v_1 with the indirect utility function that corresponds with the corner solution $\varepsilon_z = 0$. Similarly, the derivation of ε_z^{***} combines v_2 with the indirect utility function that relates to the corner solution $\varepsilon_z = 0$. The results are as follows:

$$\varepsilon_z^{**} = \upsilon_b \tag{A.4}$$

$$\varepsilon_z^{***} = \sqrt{\frac{-\nu_c - \frac{1}{2}(\nu_b)^2}{\nu_a + \frac{1}{2}}} \tag{A.5}$$

As mentioned in the text. the sign of $\varepsilon_z^* - \varepsilon_z^{**}$ determines whether the health care demand function consists of three or two pieces. The sign of $\varepsilon_z^* - \varepsilon_z^{**}$ in turn is determined by the relation between the co-payment maximum m , and a critical value for this co-payment maximum, say m^*. If $m > m^*$, health care demand is described by (4.6); if $m < m^*$, equation (4.8) describes health care demand. The expression for m^* follows from elaborating the condition $\varepsilon_z^* = \varepsilon_z^{**}$ $(= \varepsilon_z^{***})$. This yields the following:

$$m^* = \frac{1 - \varepsilon_c y}{\varepsilon_c}\left\{\sqrt{1 + \frac{(b_1 t)^2 \varepsilon_c}{\varepsilon_m}} - 1\right\} \tag{A.6}$$

Appendix B Moments of the lognormal distribution

A stochastic variable *x* is said to be log-normally distributed if its logarithm, ln *x*, has a normal distribution with parameters, say, μ and σ. The density function g(.) of *x* obeys:

$$g(x) = \frac{1}{x\sqrt{2\pi\sigma}} \exp\left(\frac{-(\ln x - \mu)^2}{2\sigma^2} \right) \tag{B.1}$$

And it follows that:

$$\Pr[a \le x \le b] = G(b) - G(a) = \int_a^b g(x)dx = F\left(\frac{\ln b - \mu}{\sigma}\right) - F\left(\frac{\ln a - \mu}{\sigma}\right) \tag{B.2}$$

where F(.) denotes the standard normal distribution function.

From equation (B.2) it also follows that the mathematical expectation of *x, E(x),* obeys:

$$E(x) = \int_0^\infty xg(x)dx = \exp\left(\mu + \tfrac{1}{2}\sigma^2\right) \tag{B.3}$$

Using this expression we calculate the (conditional) moments of *g*.
The conditional n-th moment equals:

$$\begin{aligned} E_0\left[x^n \mid a \le x \le b\right] &= \int_a^b x^n g(x)dx = \\ &\left\{ F\left(\frac{\ln b - \mu}{\sigma} - n\sigma\right) - F\left(\frac{\ln a - \mu}{\sigma} - n\sigma\right) \right\} \left(E(x)\right)^n \exp\left(\tfrac{1}{2} n(n-1)\sigma^2\right) \end{aligned} \tag{B.4}$$

The corresponding conditional expectation per capita of the group to which this expectation applies, is slightly different:

$$E\left[x^n \mid a \le x \le b\right] = \frac{\left\{ F\left(\frac{\ln b - \mu}{\sigma} - n\sigma\right) - F\left(\frac{\ln a - \mu}{\sigma} - n\sigma\right) \right\}}{\left\{ F\left(\frac{\ln b - \mu}{\sigma}\right) - F\left(\frac{\ln a - \mu}{\sigma}\right) \right\}} \left(E(x)\right)^n \exp\left(\tfrac{1}{2} n(n-1)\sigma^2\right) \tag{B.5}$$

It is the definition in (B.5) that is relevant for our model.

From (B.4) and (B.5), it follows that the conditional absolute expectation and the conditional expectation per capita equal

$$E_0\left[x \mid a \le x \le b\right] = \int_a^b x g(x) dx = \left\{ F\left(\frac{\ln b - \mu}{\sigma} - \sigma \right) - F\left(\frac{\ln a - \mu}{\sigma} - \sigma \right) \right\} E(x) \tag{B.6}$$

$$E\ \left[x \mid a \le x \le b\right] = \frac{\left\{ F\left(\frac{\ln b - \mu}{\sigma} - \sigma \right) - F\left(\frac{\ln a - \mu}{\sigma} - \sigma \right) \right\}}{\left\{ F\left(\frac{\ln b - \mu}{\sigma} \right) - F\left(\frac{\ln a - \mu}{\sigma} \right) \right\}} E(x) \tag{B.7}$$

And the second-order moments equal

$$\begin{aligned} & E_0\left[x^2 \mid a \le x \le b\right] = \int_a^b x^2 g(x) dx = \\ & \left\{ F\left(\frac{\ln b - \mu}{\sigma} - 2\sigma \right) - F\left(\frac{\ln a - \mu}{\sigma} - 2\sigma \right) \right\} (E(x))^2 \exp\left(\sigma^2\right) \end{aligned} \tag{B.8}$$

$$E\left[x^2 \mid a \le x \le b\right] = \frac{\left\{ F\left(\frac{\ln b - \mu}{\sigma} - 2\sigma \right) - F\left(\frac{\ln a - \mu}{\sigma} - 2\sigma \right) \right\}}{\left\{ F\left(\frac{\ln b - \mu}{\sigma} \right) - F\left(\frac{\ln a - \mu}{\sigma} \right) \right\}} (E(x))^2 \exp\left(\sigma^2\right) \tag{B.9}$$

The unconditional variance can be obtained from (B.3) and (B.4):

$$\begin{aligned} & \sigma_x^2 = E(x^2) - (E(x))^2 = \int_0^\infty x^2 g(x) dx - \exp\left(2\mu + \sigma^2\right) = \\ & \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right) = (E(x))^2 \left(\exp\left(\sigma^2\right) - 1\right) \end{aligned} \tag{B.10}$$

From equation (B.6), the following expression for the coefficient of variation (the ratio of standard deviation and expected value) of a log-normally distributed variable can be derived:

$$C_v = \sqrt{\left(\exp(\sigma^2) - 1\right)} \tag{B.11}$$

So, the coefficient of variation of the lognormal distribution of a variable *x* only depends on the variance of ln *x*.

References

Arrow, K.J., 1963, Uncertainty and the Welfare Economics of Medical Care, *American Economic Review* 53, 941-973.

Blomqvist, A., 1997, Optimal Non-Linear Health Insurance, *Journal of Health Economics* 16, 303-321.

Duan, N., W.G. Manning, C.N. Morris and J.P. Newhouse, 1983, A Comparison of Alternative Models for the Demand for Medical Care, *Journal of Business and Economic Statistics* 1, 115-126.

Ellis, R.P., 1986, Rational Behavior in the Presence of Coverage Ceilings and Deductibles, *RAND Journal of Economics* 17, 158-175.

Feenberg, D. and J. Skinner, 1994, The Risk and Duration of Catastrophic Health Care Expenditures, *Review of Economics and Statistics,* 633-647.

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

Feldstein, M. S., 1973, The Welfare Loss of Excess Health Insurance*, Journal of Political Economy* 81, 251-280.

Gerdtham, U.-G., J. Søgaard, B. Jönsson and F. Andersson, 1992, A Pooled Cross-Section Analysis of the Health Care Expenditures of the OECD Countries, in P. Zweifel and H.E. Frech III ,eds., *Health Economics Worldwide*, Kluwer Academic Publishers, Dordrecht, pp. 287-310.

Hausman, J.A., 1985, The Econometrics of Nonlinear Budget Sets, *Econometrica* 53, 1255-1282.

Keeler, E.B., J.P. Newhouse and C.E. Phelps, 1977, Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule under Uncertainty, *Econometrica* 45, 641-656.

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

Mirrlees, J.A., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175 - 208.

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

Pauly, M., 1968, The Economics of Moral Hazard, *American Economic Review* 53, pp. 531-537.

Phelps, C.E. and J.P. Newhouse, 1974, Coinsurance, the Price of Time, and the Demand for Medical Services, *Review of Economics and Statistics* 56, 334-342.

Raviv, A., 1979, The Design of an Optimal Insurance Policy, *American Economic Review* 69, 84-96.

Robinson, R., 2002, User Charges for Health Care, in E. Mossialos, A. Dixon, J. Figueras and J. Kutzin ,eds., *Funding Health Care: Options for Europe,* European Observatory on Health Care Systems Series, Open University Press: Buckingham, 161-183.

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

Sandmo, A., 1987, A Reinterpretation of Elasticity Formulae in Optimum Tax Theory, *Economica* 54, 89-96.

Van Vliet, R.C.J.A., 1998, Schatting verzekeringseffecten op basis van de CBS Gezondheidsenquête ten behoeve van Ramingsmodel Zorgsector, iBMG, Erasmus University Rotterdam, *in Dutch*.

Van Vliet, R.C.J.A. and H.G. van der Burg, 1996, Verdelingsfuncties voor Kosten van Zorgvoorzieningen, iBMG, Erasmus University Rotterdam *,in Dutch*.

Wedig, G.J., 1988, Health Status and the Demand for Health – Results on Price Elasticities, *Journal of Health Economics* 7, 151-163.

Zweifel, P. and W.G. Manning, 2000, Moral Hazard and Consumer Incentives in Health Care, in A.J. Culyer and J.P. Newhouse ,eds., *Handbook of Health Economics*, Elsevier, North-Holland, pp. 409-459.