Optimal cost sharing to avoid risk selection in health insurance markets

Mathias Kifmann   Normann Lorenz*†

December 2003

Abstract

In the absence of a perfect risk adjustment scheme, cost sharing can reduce risk selection in community-rated health insurance markets. However, cost sharing also reduces incentives for efficiency. In this paper, we develop a model in which insurers determine the cost efficiency of health care. They have incentives for risk selection because there are two risk types which differ in their expected health care costs. We derive the optimal cost sharing function and show that costs should be shared where the cumulative cost reimbursement for high risk types compared to low risk types is comparatively large and where there is a low concentration of individuals. Using individual health cost data from a Swiss health insurer, we show that an optimal cost sharing scheme should reimburse costs only up to a limit.

JEL-**classification:** H42, I18.

Keywords: health insurance, risk selection, risk adjustment.

---

*Universität Konstanz, Fach D 136, 78457 Konstanz, Germany; Email: Mathias.Kifmann@uni-konstanz.de, Normann.Lorenz@uni-konstanz.de.

†We thank Konstantin Beck from CSS Insurance, Switzerland, and Andreas Werblow for making the data set available to us. We are grateful for helpful comments and suggestions by Friedrich Breyer, Stefan Felder, Thomas McGuire, Thomas Weber, and participants at the International Health Economics Association 4th World Congress in San Francisco.

1 Introduction

Risk selection is a major concern in community-rated health insurance markets. As Pauly (1984) has pointed out, insurers have an incentive to discriminate against high risks and to attract low risks in such markets since they are not allowed to charge risk-based premiums. To avoid risk selection, regulators frequently impose open enrollment and define standardized benefit packages. However, these measures are likely to be insufficient. In recent years, an important research topic has therefore been the design of incentive systems which discourage risk selection.[1]

Most of the literature has focused on risk adjustment schemes which reallocate funds among health insurers according to observable characteristics of their insured. In practice, risk adjustment cells are defined on these characteristics. The transfer payment for an individual is then determined by the average cost of all insured in the respective cell.[2] However, it remains unclear whether risk adjustment schemes can sufficiently reduce risk selection. The main problem is the availability of data. Usually, only few characteristics such as age and gender can easily be obtained. Further indicators, in particular diagnostic information, are only available at a considerable cost.[3] Even if risk adjustment schemes are considerably improved, risk selection may still be highly profitable (Newhouse (1994)).

In the absence of perfect risk adjustment, cost sharing can be an important method to counter risk selection. Although cost sharing naturally leads to efficiency losses, this may be a price worth paying to reduce risk selection. Although the possible benefits of cost sharing are generally recognized, there has been little theoretical work on the characteristics of an optimal cost sharing function. Usually cost sharing is regarded as a "mandatory reinsurance program" scheme.[4] This analogy suggests that optimal cost sharing is similar to an optimal insurance con-

---

[1] For a survey, see the chapter on risk adjustment in competitive health plan markets by van de Ven and Ellis (2000) in the Handbook of Health Economics.

[2] Recent contributions by Glazer and McGuire (2000, 2002) and Frank, Glazer, and McGuire (2000) show that this approach is not optimal if observable characteristics are only imperfect signals for an individual's health status. They demonstrate that a risk adjustment scheme which takes this into account can be much more effective in avoiding risk selection.

[3] In a recent paper, Marchand, Sato, and Schokkaert (2003) show that prior expenditure can also be a useful risk adjuster to avoid risk selection.

[4] See van de Ven and Ellis (2000, p. 818).

tract. Here Arrow (1974) has shown that a deductible and full coverage above the deductible is optimal if the insurer is risk neutral and his cost function is linear. Raviv (1979) extended his result to risk averse insurers and strictly convex cost functions. In this case, co-insurance above the deductible characterizes the optimal insurance contract. For optimal cost sharing this implies that costs should only be reimbursed above a threshold. This *outlier risk sharing* (van de Ven and Ellis (2000)) is used in practice. In Germany, for example, 60% of individual health care costs which exceed €20,450 are reimbursed.

In this paper, we take a different perspective on cost sharing. We focus on the *incentive problem* of minimizing risk selection without compromising too heavily on efficiency and consider explicitly the incentives for risk selection and cost efficiency for insurers. In particular, we assume that insurers influence the cost of health care by organizing the delivery in a more efficient way or by negotiating lower prices with providers. Incentives for risk selection arise because two risk types differ in their expected health care costs. The regulator cannot observe the risk types. We assume, however, that he knows how health care costs of each type are distributed. To determine the optimal cost sharing function, we minimize the difference in expected costs between the risk types for a given increase in total costs.

We find that the optimal cost sharing function depends on two factors. First, costs should be shared where the cumulative cost reimbursement for high risk types compared to low risk types is comparatively large. This factor depends on the difference in the distribution functions of the risk types. Second, cost sharing should be avoided where there is a large concentration of individuals and therefore a high efficiency cost. This factor relies on the joint density function of health care costs. This result demonstrates that outlier risk sharing can be inefficient if costs are shared at the "wrong" part of the distribution of costs, i.e. where the comparative effect on high risk types is low or where there is a large concentration of individuals which leads to high efficiency costs. Our cost sharing function explicitly takes into account these factors.

We use health cost data from a Swiss health insurer to illustrate how the optimal cost sharing formula can look like in practice. It turns out that costs should generally be shared only up to a limit which is the exact opposite of outlier risk sharing. Comparing optimal cost sharing with outlier risk sharing we find that our approach is much more effective. The mean decrease in the average cost difference between the two groups is almost three times larger if optimal cost sharing is used instead of outlier risk sharing. In addition, outlier risk sharing may actually increase the incentives for risk selection because it can increase the difference in average costs between the groups. These results show that the theory of optimal insurance does not carry over to optimal cost sharing to avoid risk selection. The reason is that an optimal insurance contract is designed to share an uncertain payoff of a risk-averse party but cannot deal adequately with the incentive problem of minimizing risk selection without compromising too strongly on cost efficiency. For this reason, we prefer to use the term "cost sharing" instead of "risk sharing".

Our optimal cost sharing function applies to all individuals. In the literature, other forms of cost sharing have been put forward which are limited to a specific group.[5] Van de Ven and van Vliet (1992) propose *risk sharing for high risks* which allows an insurer to designate a specified percentage of his insured for which all health care costs will be reimbursed. *Risk sharing for high costs* is considered by van Barneveld, Lamers, van Vliet, and van de Ven (2001). Under this scheme, all health care costs of a predetermined number of individuals with the highest costs are paid by the regulator. In an empirical study, van Barneveld, Lamers, van Vliet, and van de Ven (2001) compare these forms of risk sharing to outlier risk sharing and proportional risk sharing which reimburses a fixed percentage of all costs. They find that *risk sharing for high risks* as well as *risk sharing for high costs* is superior in reducing incentives for risk selection to outlier risk sharing and proportional risk sharing. These forms of risk sharing are more effective in reimbursing only the costs of high risk types without sharing the costs of low risk types.

Since we assumed that the cost sharing formula applies to all individuals, we cannot say whether our approach is superior to these selective forms of cost sharing. However, our approach is explicitly designed to share only the costs of high risk types and is superior to outlier and proportional risk sharing. In future comparisons of different cost sharing approaches it would therefore be interesting to

---

[5] A further possibility is to make cost sharing dependent on a medical condition; see van de Ven and Ellis (2000, p. 822).

compare risk sharing for high risks or high costs to our optimal cost sharing approach.

The paper is structured as follow. In Section 2, we present the model. The optimal cost sharing formula is derived and discussed in Section 3. In Section 4, we calculate optimal cost sharing formulas based on data from a Swiss health insurer. Section 5 summarizes the results and concludes.

2 The model

We analyze a health insurance market in which the regulator wants to make medical services available to all individuals at a uniform price. He imposes community rating and requires all insurers to offer a standardized health insurance package. Regulation also includes open enrollment, i.e. insurers must accept all individuals applying for insurance.

Health insurers organize the delivery of medical services. By choosing higher effort e health insurers can organize the delivery in a more efficient way or negotiate lower prices with providers.[6] Choosing a higher effort will therefore decrease costs to treat an illness but also the utility of an insurer. For the disutility of effort $v(e)$ we assume $v'(e) > 0$ and $v''(e) \geq 0$. Costs to treat a patient depend on the effort level e and on the severity m of the patient's illness where $0 \leq m \leq M$. We assume

$$C(e,m) = c(e)m, \quad c(e) > 0, c'(e) < 0, c''(e) > 0, \tag{1}$$

i.e. costs are proportional to m and organizational effort is subject to decreasing returns to effort. The effort level chosen by an insurer when there is no cost sharing is labelled $\hat{e}$. We normalize $c(\hat{e}) = 1$ so that $C = m$ in the absence of cost sharing.

An individual can be a high risk h or a low risk l. Expected costs of the high risk type are larger than expected costs of the low risk type. The proportion of l-types is θ. For each risk type $i = l,h$, severity m is distributed according to the distribution function $F_i(m)$. Since a substantial fraction of insured usually does

---

[6] See Marchand, Sato, and Schokkaert (2003) for a similar approach.

not use any health services during a certain period we allow for $F_i(0) > 0$. We assume the distribution function to be continuously differentiable for all $m \geq 0$ and label the respective density function $f_i(m)$.[7] For $m > 0$ we have

$$F_i(m) = F_i(0) + \int_0^m f_i(s)\,\mathrm{d}s. \tag{2}$$

Expected costs of each risk type correspond to

$$\mathrm{E}_i[C(e,m)] = \int_0^M c(e) m f_i(m)\,\mathrm{d}m \tag{3}$$

with $\mathrm{E}_h[C(e,m)] > \mathrm{E}_l[C(e,m)]$.

We assume that the regulator cannot identify the risk type and therefore is not able to implement a perfect risk adjustment scheme. Neither can he observe e nor m. However, he knows the distribution functions $F_i(m)$ for each risk type. For example, he may infer the distribution functions from a representative sample with information about the risk type.[8] His objective is to find a balance between incentives for risk selection and efficiency by sharing costs with insurers. For an individual with cost $C(e,m)$ he reimburses $r(C(e,m))$. With respect to $r(C)$, we impose two restrictions:

1. $r'(C) \leq 1$ – no incentives for cost-inflation

   This restriction guarantees that the insurer cannot increase his profits by inflating costs.

2. $r'(C) \geq 0$ – no incentives for cost-deflation

   If $r(C)$ is non-decreasing in C, then hiding costs cannot lead to higher profits for the insurer.

---

[7] To be more precise, we assume the distribution function $F_i(m)$ to be continuously differentiable for all $m > 0$ and $\lim_{m\to 0^+} F_i'(m)$ to exist; accordingly to simplify the notation by $f_i(0)$ we mean $\lim_{m\to 0^+} f_i(m)$.

[8] See section 4 for an illustration of this procedure.

We also assume that the cost sharing scheme has a balanced budget, i.e.

$$r(C(e,0))[\theta F_l(0)+(1-\theta)F_h(0)]+\int_0^M r(C(e,m))[\theta f_l(m)+(1-\theta)f_h(m)]=0. \quad (4)$$

This means that the cost sharing scheme is self-financing, which implies that $r(C(e,0))$ must be negative if costs are reimbursed.[9]

Health insurers know to which group a particular insured belongs. Because health insurance premiums are community-rated, health insurers try to risk select if average costs of the two groups differ. They can do so by imposing barriers for high risk individuals and trying to attract low risk individuals. For example, they may process applications of high risks only slowly. Low risks, on the other hand, may be captured by special package deals.[10] Taking into account cost sharing by the regulator, the difference in expected cost (DEC) between the two risk types

$$\begin{aligned} DEC &= \mathrm{E}_h[C(e,m)-r(C(e,m))]-\mathrm{E}_l[C(e,m)-r(C(e,m))] \\ &= -r(C(e,0))F_h(0)+\int_0^M [C(e,m)-r(C(e,m))]\,f_h(m)\,\mathrm{d}m \qquad (5) \\ &\quad -\Big(-r(C(e,0))F_l(0)+\int_0^M [C(e,m)-r(C(e,m))]\,f_l(m)\,\mathrm{d}m\Big) \end{aligned}$$

captures the extra profit for an insurer when he insurers a low risk type instead of a high risk type. We assume that the incentives to risk select are higher for an insurer, the larger this difference.

Risk selection is a zero sum game between insurers in which every insurer spends resources to attract a favorable selection of individuals. We do not explicitly model this contest but in any symmetric equilibrium, each insurer ends up with a representative share of the two risk groups. We assume that risk selection leads to a loss of resources which is increasing in the difference in expected cost between the two risk types.

---

[9] In practice, cost sharing schemes are frequently financed by a uniform flat rate and define a nonnegative cost sharing function. In such a framework, $-r(C(e,0))$ corresponds to the uniform flat rate and $r(C(e,m))-r(C(e,0))$ equals the cost reimbursement.

[10] We therefore focus on what Glazer and McGuire (2002, p. 154) have termed the *access problem* and do not analyze the incentives of health insurers to distort the mix of the quality of health care.

With the design of the cost sharing function $r(C(e,m))$, the regulator can influence the difference in expected cost and therefore the loss of resources due to risk selection. However, cost sharing will also reduce effort from the first-best level $\hat{e}$ to $\tilde{e}$ and therefore lead to higher average costs

$$\begin{aligned} AC(\tilde{e}) &= \theta \mathrm{E}_l \left[C\left(\tilde{e}, m\right)\right] + (1-\theta)\mathrm{E}_h \left[C\left(\tilde{e}, m\right)\right] \\ &\qquad > AC(\hat{e}) = \theta \mathrm{E}_l \left[C\left(\hat{e}, m\right)\right] + (1-\theta)\mathrm{E}_h \left[C\left(\hat{e}, m\right)\right] \end{aligned} \tag{6}$$

The problem of the regulator is to choose the cost sharing function $r(C)$ such that the difference in expected cost is reduced without increasing average costs too much.

The sequence of events is as follows:

1. The regulator announces the reimbursement function $r(C)$.
2. Insurers expend resources to risk select.
3. Individuals choose insurers, each insurer ends up with a representative share of the two risk groups.
4. Insurers select organizational effort e.
5. The severity m and costs $C(m,e)$ are determined.
6. The regulator reimburses $r(C(m,e))$.

Our solution approach to determine the optimal function $r(C)$ is as follows. We assume that the regulator is willing to tolerate an increase in average costs by a factor $\tilde{x} > 1$. This assumption defines a level of organizational effort $\tilde{e}$ at stage 4. We solve problem (P1)

$$\begin{aligned} \min_{r(C)} DEC \qquad \text{subject to} \qquad & AC(\tilde{e}) = \tilde{x} AC(\hat{e}) \\ & \text{balanced budget condition (4)} \\ & 0 \leq r'(C) \leq 1 \end{aligned} \tag{P1}$$

This yields the optimal cost sharing function $r(C)$ which minimizes the difference in expected costs for insurers and therefore incentives to risk select at stage 2.

The advantage of this approach is twofold. First, we avoid assumptions about the size of loss of resources due to risk selection. In addition, we do not need to specify the trade-off the regulator is willing to make between avoiding risk selection and preserving incentives for cost efficiency. This information is not necessary to characterize the main properties of the optimal cost sharing function.

3 The optimal cost sharing formula

To solve problem (P1) we proceed in four steps:

1. We define the effort level $\tilde{e}$ which is associated with a given increase in average costs by a factor $\tilde{x} > 1$.
2. We determine the incentive constraint which guarantees that insurers choose effort level $\tilde{e}$.
3. We reformulate problem (P1) as the optimal control problem (P2) with costs C as the integration variable.
4. We solve the optimal control problem and characterize the optimal cost sharing function $r(C)$.

Step 1: Our assumption is that the regulator is willing to tolerate an increase in average cost by a factor $\tilde{x}$ compared to a situation of no cost sharing. The corresponding effort level $\tilde{e}$ is defined by the condition

$$\begin{aligned} AC(\tilde{e}) &= \theta \mathrm{E}_l\left[c(\tilde{e})m\right] + (1-\theta)\mathrm{E}_h\left[c(\tilde{e})m\right] \\ &= \tilde{x} AC(\hat{e}) = \tilde{x}\Big(\theta \mathrm{E}_l\left[c(\hat{e})m\right] + (1-\theta)\mathrm{E}_h\left[c(\hat{e})m\right]\Big) \end{aligned} \tag{7}$$

where $\hat{e}$ is the first-best effort level in absence of cost sharing. Since we normalize $c(e)$ such that $c(\hat{e}) = 1$ in a situation without cost sharing, condition (7) simplifies to $c(\tilde{e}) = \tilde{x}$ which implies

$$\tilde{e} = c^{-1}(\tilde{x}). \tag{8}$$

Step 2: When insurers select effort e at stage 4 and face a cost sharing function $r(C)$, their optimization problem is

$$\max_{e} P - \theta \mathrm{E}_h \left[c(e)m - r(c(e)m)\right] - (1-\theta)\mathrm{E}_l \left[c(e)m - r(c(e)m)\right] - v(e)$$

where P is the insurance premium net of expenditure on risk selection. The first-order condition is

$$-\theta \mathrm{E}_h \left[c'(e)m - r'(c(e)m)c'(e)m\right] - (1-\theta)\mathrm{E}_l \left[c'(e)m - r'(c(e)m)c'(e)m\right] - v'(e)$$

which corresponds to

$$-\int_0^M [c'(e)\,m - r'(c(e)m)c'(e)m][\theta f_l(m) + (1-\theta) f_h(m)]\,\mathrm{d}m - v'(e) = 0. \tag{9}$$

Rearranging terms yields

$$\int_0^M r'(c(e)m)mg(m)\,\mathrm{d}m = k(e) \quad \text{with} \quad k(e) \equiv \int_0^M mg(m)\,\mathrm{d}m + \frac{v'(e)}{c'(e)} \tag{10}$$

where $g(m) \equiv \theta f_l(m) + (1-\theta) f_h(m)$ is the average density function. A sufficient condition for the corresponding effort level to yield a global profit-maximum is that the cost sharing function $r(C)$ is concave.[11] From condition (10), it follows that the cost sharing function must satisfy

$$\int_0^M r'(c(\tilde{e})m)mg(m)\,\mathrm{d}m = k(\tilde{e}) \tag{11}$$

if insurers are to choose effort level $\tilde{e}$. Condition (11) therefore defines the incentive constraint which guarantees that average costs increase by the factor $\tilde{x}$.

---

[11] If $r(C)$ is a concave function, then the left hand side of (10) is non-decreasing in e. Since the function $k(e)$ is a strictly decreasing function of e, the first-order condition must therefore characterize a global optimum. If $r(C)$ is not concave, then it needs to be checked whether equation (10) guarantees an optimum.

Step 3: To derive the optimal cost sharing function $r(C)$ it is convenient to express our problem such that C is the integration variable. We therefore transform the distribution functions $F_i(m)$ and the density functions $f_i(m)$ into functions of C. From $C = c(\tilde{e})m$ it follows that $m = C/c(\tilde{e})$. The distribution functions in terms of C are therefore given by $\tilde{F}_i(C) = F_i(C/c(\tilde{e}))$. Differentiating with respect to C yields the corresponding density functions $\tilde{f}_i(C) = f_i(C/c(\tilde{e}))/c(\tilde{e})$. The support of C is given by $[0, c(\tilde{e})M] = [0, \overline{C}]$.

Using this transformation in the definition of the difference in expected cost (equation (5)), we obtain

$$\begin{aligned} DEC &= -r(0)\tilde{F}_h(0) + \int_0^{\overline{C}} [C - r(C)]\, \tilde{f}_h(C)\, \mathrm{d}C \\ &\quad - \Big(-r(0)\tilde{F}_l(0) + \int_0^{\overline{C}} [C - r(C)]\, \tilde{f}_l(C)\, \mathrm{d}C \Big). \end{aligned} \tag{12}$$

Noting that $\int_0^{\overline{C}} C[\tilde{f}_h(C) - \tilde{f}_l(C)]\mathrm{d}C$ is a constant, the regulator's problem (P1) therefore is equivalent to problem (P2)[12]

$$\min_{r(C)} \quad r(0)[\tilde{F}_l(C) - \tilde{F}_h(C)] + \int_0^{\overline{C}} r(C)[\tilde{f}_l(C) - \tilde{f}_h(C)]\mathrm{d}C$$

subject to

$$\begin{aligned} &\int_0^{\overline{C}} r'(C) C \tilde{g}(C)\, \mathrm{d}C = c(\tilde{e})k(\tilde{e}) \qquad \text{(P2)} \\ &r(0)\tilde{G}(0) + \int_0^{\overline{C}} r(C)\tilde{g}(C)\, \mathrm{d}C = 0 \\ &0 \le r'(C) \le 1 \\ &r(0), r(\overline{C}) \text{ free} \end{aligned}$$

where $\tilde{g}(C) = \theta \tilde{f}_l(C) + (1-\theta)\tilde{f}_h(C)$ and $\tilde{G}(C) = \theta\tilde{F}_l(C) + (1-\theta)\tilde{F}_h(C)$. The first constraint is the transformed incentive constraint (11) which ensures that average

---

[12] We assume that the difference in expected costs remains positive. If in the optimum this difference were negative, then cost sharing would clearly be too high and a lower value of $\tilde{x}$ should be chosen.

costs increase only by the factor $\tilde{x}$. The second constraint corresponds to the zero-budget constraint (4). The third constraint ensures that there are neither incentives for cost-inflation nor cost-deflation. Finally, the last constraint expresses that there are no restrictions with respect to the endpoints of $r(C)$.

Step 4: Problem (P2) is an isoperimetric dynamic optimization problem due to the equality integral constraints.[13] It is not possible to set up the Hamiltonian and to apply the maximum principle since we allow for $F_i(0) > 0$. In the Appendix, we therefore formulate the Lagrangian function for the full problem. There, we derive the following result

Proposition 1: *The slope of the optimal cost sharing function is characterized by*

$$(1-\tilde{F}_h(C)) - (1-\tilde{F}_l(C)) - |\bar{\eta}| C\tilde{g}(C) \begin{cases} > 0 \Rightarrow & r'(C) = 1 \\ = 0 \Rightarrow & 0 \leq r'(C) \leq 1 \\ < 0 \Rightarrow & r'(C) = 0 \end{cases} \quad (13)$$

where $\bar{\eta} < 0$ *is the constant Lagrange-multiplier associated with the incentive constraint. Condition* (13) *and the zero-budget constraint*

$$r(0)\tilde{G}(0) + \int_0^{\bar{C}} r(C)\tilde{g}(C)\, dC = 0$$

determine $r(0)$.

To interpret this result, first note that the optimal cost sharing formula $r(C)$ has a slope of either zero or one unless by chance we have $(1-\tilde{F}_h(C)) - (1-\tilde{F}_l(C)) = \bar{\eta} C\tilde{g}(C)$. Furthermore, condition (13) can be decomposed into two terms with a natural interpretation:

---

[13] See Chiang (1992, p. 280) and Kamien and Schwartz (1991, p. 228).

1. *The anti-selection term* $(1-\tilde{F}_h(C))-(1-\tilde{F}_h(C))$

   A large difference $(1-\tilde{F}_h(C))-(1-\tilde{F}_h(C))$ tends to favor cost sharing. To explain this effect, it is important to note that $r'(C)=1$ increases cost sharing for all individuals with costs *above* C. This follows from the restriction $r'(.)\geq 0$. Since $1-\tilde{F}_i(C)$ denotes the share of i-types with costs higher than C, $(1-\tilde{F}_h(C))>(1-\tilde{F}_l(C))$ implies that there are relatively more h-types with costs above C than l-types. Increasing cost sharing at C therefore reimburses costs more for h-types than for l-types. This implies that the difference in expected costs must fall.

2. *The cost efficiency term* $|\bar{\eta}|C\tilde{g}(C)$

   A large value of $|\bar{\eta}|C\tilde{g}(C)$ calls for no cost sharing. This is because $C\tilde{g}(C)$ corresponds to the share of average expected costs at C. If this share is large, then cost sharing at C tends to have a large negative impact on the incentives for efficiency and therefore calls for no cost sharing. This effect is increasing in $|\bar{\eta}|$, the Lagrange-multiplier which captures the importance of incentives for cost efficiency. A higher $\tilde{e}$, i.e. higher incentives for cost efficiency increases $|\bar{\eta}|$. For a given value of the anti-selection term, a larger value of $|\bar{\eta}|$ therefore implies less cost sharing.

The optimal cost sharing functions therefore considers for every cost level C whether the reduction in incentives for risk-selection outweigh the efficiency costs of cost sharing. The restriction that costs are only allowed to increase by a certain percentage is reflected in the Lagrange-multiplier $\bar{\eta}$ which defines the importance of cost-efficiency.

An example

A simple example is useful to illustrate our result. Suppose that there are two groups of equal size ($\theta=0.5$) and that $0\leq m\leq 10$. For the distribution functions we assume $f_l(m)=0.2-0.02m$ and $f_h(m)=0.02m$ with $F_l(0)=0$ and $F_h(0)=0$. The cost and the disutility function take the form $v(e)=e$ and $c(e)=5/e$.[14]

---

[14] Note that the optimal cost sharing function for a given increase in average costs is independent of specific assumptions about $c(e)$ and $v(e)$. As long as the second-order condition of the health insurer's problem is fulfilled, all functions $c(e)$ and $v(e)$ will yield the same optimal cost sharing function.

Without cost sharing, we have optimal effort $\hat{e} = 5$ and $c(\hat{e}) = 1$. The corresponding difference in expected costs is

$$\begin{aligned} \mathrm{E}_h[C(\hat{e})] - \mathrm{E}_l[C(\hat{e})] &= \int_0^1 c(\hat{e}) m f_h(m) dm - \int_0^1 c(\hat{e}) m f_l(m) dm \\ &= 6.67 - 3.33 = 3.33. \end{aligned}$$

We assume that the regulator wants to induce a level of effort $\tilde{e}$ such that average costs increase by 5%. Thus, $\tilde{e}$ is defined by $c(\tilde{e}) = 1.05$ by condition (8) which implies $\tilde{e} = 4.762$ and yields

$$\begin{aligned} (1-\tilde{F}_h(C)) - (1-\tilde{F}_l(C)) &= 0.1905C - 0.0181C^2 \\ \tilde{g}(C) &= 0.095238. \end{aligned}$$

Solving the optimization problem, we find for the Lagrange-multiplier associated with the incentive constraint $\overline{\eta} = -1.3902$. Using condition (13), i.e.

$$(1-\tilde{F}_h(C)) - (1-\tilde{F}_l(C)) - |\overline{\eta}| C \tilde{g}(C) \begin{cases} > 0 & \Rightarrow \quad r'(C) = 1 \\ = 0 & \Rightarrow \quad 0 \le r'(C) \le 1 \\ < 0 & \Rightarrow \quad r'(C) = 0 \end{cases}$$

we can infer the optimal cost sharing function: Figure 1 shows the anti-selection term $(1-\tilde{F}_h(C)) - (1-\tilde{F}_l(C))$ and the cost efficiency term $|\overline{\eta}| C \tilde{g}(C)$ depending on C. The functions intersect at $C = 3.20$. Where the anti-selection term is larger than the cost efficiency term, marginal cost sharing equals one, where it is below there is no marginal cost sharing. Thus, we obtain

$$r'(C) = \begin{cases} 1 & \text{for} \quad C \le 3.20 \\ 0 & \text{for} \quad C > 3.20 \end{cases}.$$

The balanced-budget condition requires $r(0) = -2.71$. This yields the optimal cost sharing function

$$r(C) = \begin{cases} -2.71 + C & \text{for } C \le 3.20 \\ 0.49 & \text{for } C > 3.20 \end{cases}$$

which is shown in Figure 2.

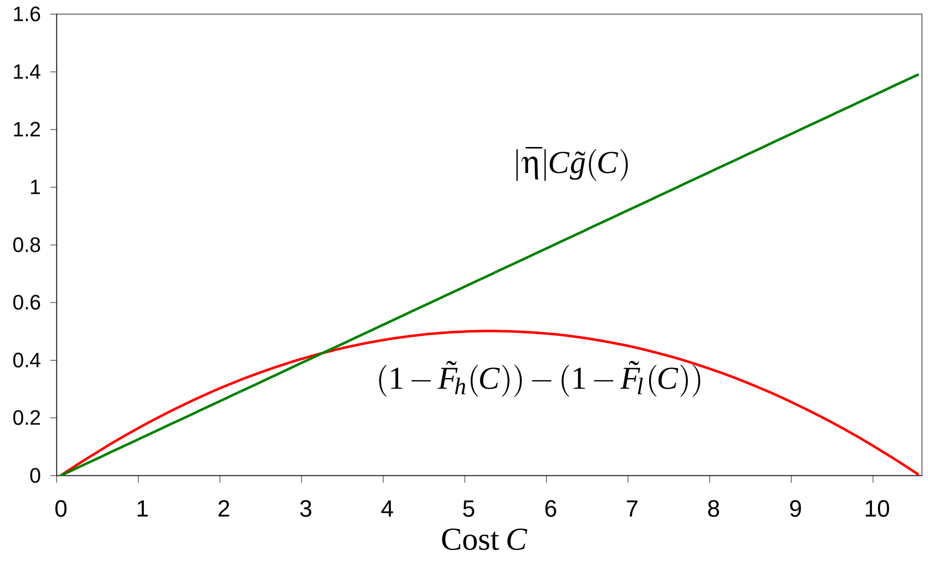

Figure 1: $(1-\tilde{F}_h(C))-(1-\tilde{F}_l(C))$ and $|\bar{\eta}|C\tilde{g}(C)$

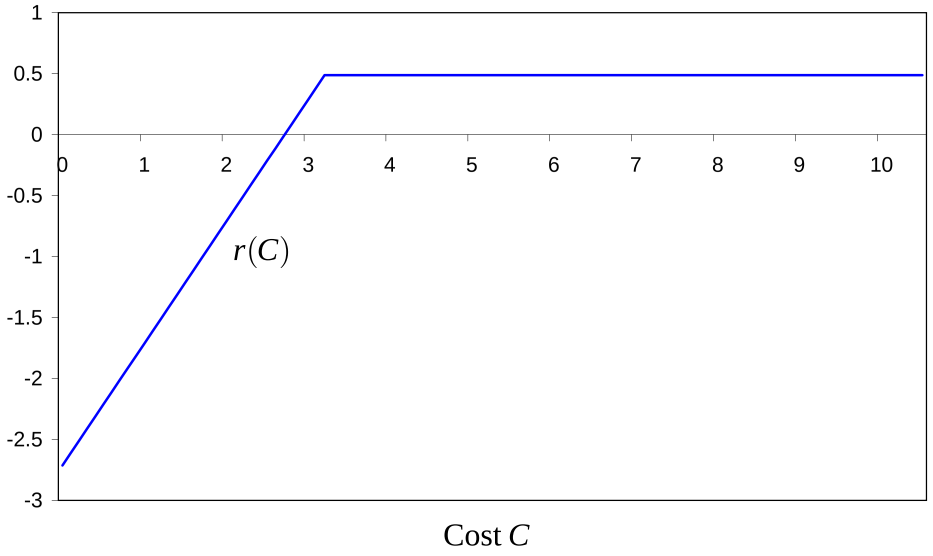

Figure 2: The optimal cost sharing function

Expected cost for the optimal cost sharing function are 6.61 and 3.89. The resulting difference in expected cost is 2.72. Thus, a 5% increase in average cost allows to reduce the difference in expected cost by $(3.33-2.72)/3.33 = 18.3\%$.

To demonstrate the efficiency of our approach, we compare our result with the difference in expected cost under outlier risk sharing. We assume that 60% of all cost above a certain threshold are reimbursed. For a 5% increase in average cost, this threshold equals 9.65. Since cost sharing has to be financed by insurers we also have a uniform rate for each insured. In this example this turns out to be -0.02. Therefore the cost sharing function with outlier risk sharing is

$$r(C) = \begin{cases} -0.02 & \text{for} \quad C < 9.65 \\ -0.02 + 0.6(C-9.65) & \text{for} \quad C \geq 9.65 \end{cases} .$$

Expected cost are 6.98 and 3.52 and the difference in expected cost is 3.46. Compared to no cost sharing, the difference in expected costs therefore *increases* by $(3.46-3.33)/3.33 = 3.8\%$. Thus, outlier risk sharing is counter-productive in this example because it leads to efficiency losses and increases the incentives for risk selection. This can be explained as follows. In the absence of cost sharing, a 5% increase in total costs also increases the difference in expected costs by 5% from 3.33 to 3.5. Outlier risk sharing can only reduce the difference in expected cost by 1.1% from 3.5 to 3.46. The inflation in the cost difference due to a general increase in costs therefore dominates the reduction in the cost difference by outlier risk sharing.

Our example shows that outlier risk sharing does not need to be optimal. Instead the optimal formula may be characterized by cost sharing up to a limit. In addition, the example demonstrates that outlier risk sharing may actually be counterproductive if the increase in the difference in average costs due to cost inflation dominates the reduction in the difference in average cost for a given effort level. In the next section, we show that these results can also arise if we apply our formula to actual health cost data.

4 An empirical illustration

The empirical analysis is based on administrative data provided by a Swiss health insurer. The data set includes information on individual costs, hospitalization, number of months insured, death and extent of coinsurance for the years 1997 to 1999 with 475,506 observations. We used the observations of 104,420 adult individuals being insured in the years 1998 and 1999. Their average health care expenditure was 3250 Swiss Franks (CHF) in 1999. We created a variable indicating to which of the 30 age-gender-cells of the Swiss risk adjustment scheme each insured belongs. Information on each group is given in Table 1 (see page 22).

To illustrate possible shapes of the optimal cost sharing function, we apply our method to each age-gender-cell. Our risk selection hypothesis is that health insurers can observe whether an individual was treated in a hospital in 1998. The group h is therefore given by those treated in a hospital in 1998. The l-types are the remaining individuals. The regulator is not able to obtain information on hospitalization or does not want to use it in a risk adjustment scheme for incentive reasons. However, on the basis of the data set which compromises a sample of the population it is possible to infer the distribution functions of the two groups and to apply our method.[15] We proceeded in two steps:

- As in our example we assume $v(e) = e$ and $c(e) = \beta/e$ where the constant β is chosen such that $c(\hat{e}) = 1$. The costs from our data set correspond to the costs if insurers choose the first-best effort level since there is no cost sharing in Switzerland. Thus, we obtain $C^{\text{act.}} = c(\hat{e})m = m$ and use actual costs $C^{\text{act.}}$ to estimate the distribution functions $F_i(m)$ for each group.
- Second, we derive the distribution functions for an $\tilde{x}$% increase in costs and apply our method as in the example.

We estimate the distribution function of actual costs $C^{\text{act.}}$ in 1999 for the two groups nonparametrically by a kernel density estimation.[16] Figure 3 shows the

---

[15] For a real world application, our data set would need to be adjusted if it is not representative.

[16] Since there was a considerable share of observations with zero costs, we set $F_i(0)$ equal to this share and determined only $F_i(C^{\text{act.}})$ with $C^{\text{act.}} > 0$ with the kernel density estimator. We chose a constant bandwidth for the kernel. We found that $f_i(C^{\text{act.}}) = 0$ for a number of intervals

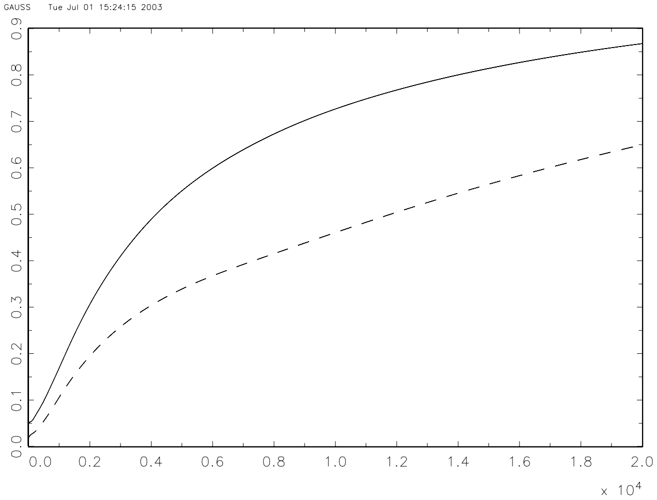

Figure 3: Distribution functions for the two groups, cell M14

distribution functions for the two groups for one age-gender-bracket for $C^{\text{act.}} \leq 20,000$ CHF. In this case, $F_l(C) > F_h(C)$ for all values of C.

Applying our method, we find that the optimal cost sharing function takes one of three types:[17]

1. The first type has marginal cost sharing equal to one starting at $C = 0$ up to a threshold which is between 8,000 and 18,000 CHF, with an interval of zero marginal cost sharing of about 2,000 to 5,000 CHF. There is no additional cost sharing above this threshold. An example is shown in Figure 4.
2. The second type has no marginal cost sharing for cost below a threshold of about 20,000 and above a threshhold of about 40,000 CHF. Between 20,000 and 40,000 CHF there is marginal cost sharing equal to one. Above this threshold, there is no more cost sharing. Figure 5 shows a graph of this type.

---

for $C^{\text{act.}} > 10,000$ CHF. This artificially improved our results because $r'(C) = 1$ does not reduce incentives for efficiency at all whenever $f_i(C) = 0$ for both groups. Therefore we transformed the data using a concave function. With the function $\ln(C^{\text{act.}})$ we did not get any intervals with $f_i(C^{\text{act.}}) = 0$. From the estimated distribution function $\hat{F}_i(\ln(C^{\text{act.}}))$ we derive the distribution function $F_i(C^{\text{act.}})$ and the density function $f_i(C^{\text{act.}})$. We also performed kernel estimation with variable bandwidths as proposed by Silverman (1986). Because we considered the differences in the bandwidth too large (the largest bandwidth was about 100,000 times as large as the smallest one) and for reasons given by Terrell and Scott (1992) we did not use this procedure.

[17] We checked the health insurers' second-order conditions for each cost sharing function. They were always satisfied.

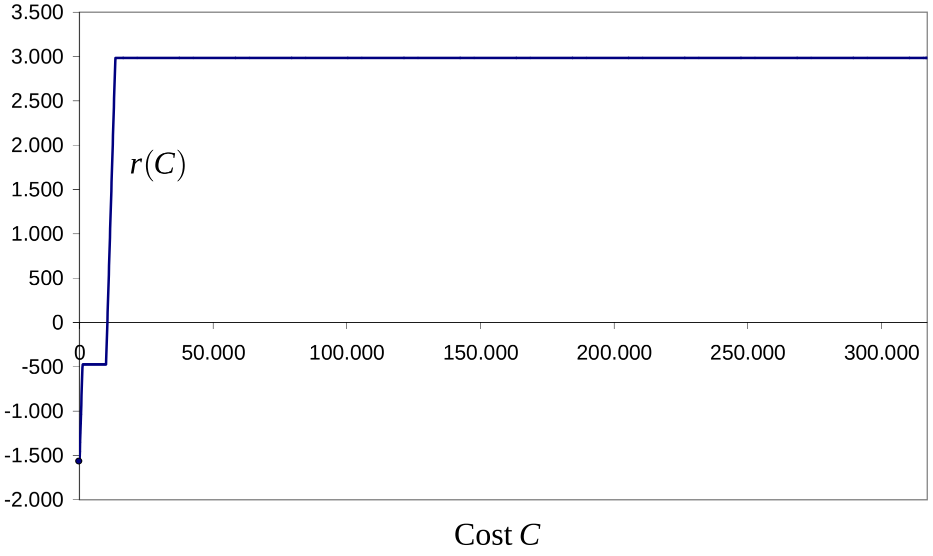

Figure 4: Optimal $r(C)$-function type 1, cell M13

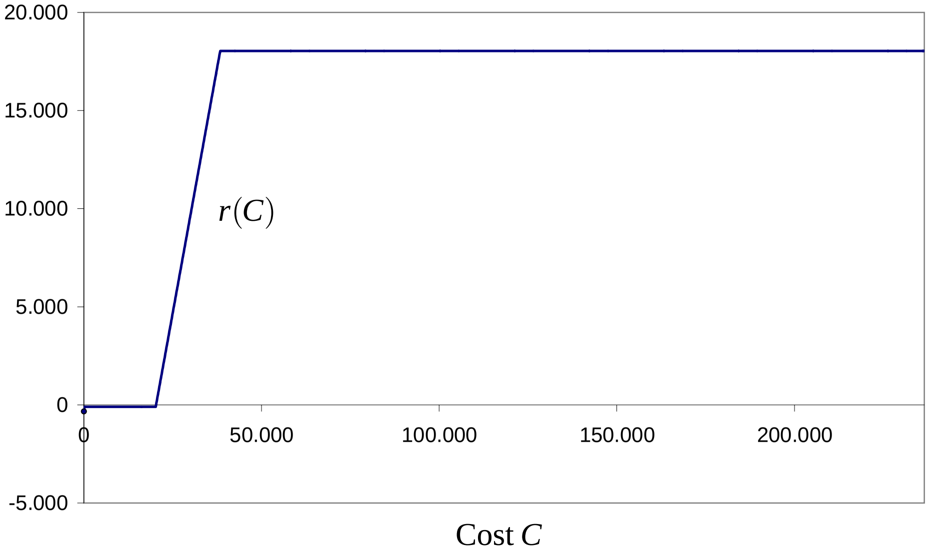

Figure 5: Optimal $r(C)$-function type 2, cell F3

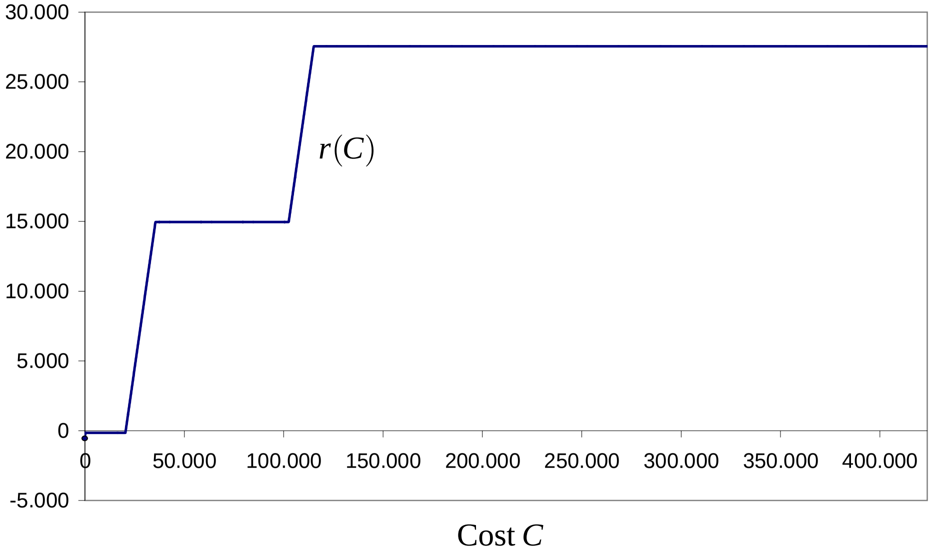

Figure 6: Optimal $r(C)$-function type 3, cell F7

3. The third type looks like the second but the threshold above which there is no additional cost sharing is much higher, mostly between 100,000 and 150,000 CHF, and with sometimes one or two intervals with no cost sharing. An example is shown in Figure 6.

For the 30 age-gender cells, 12 of the optimal cost sharing were of type one, 10 of type two and 8 of type three. Table 1 and Figures 7 to 9 give an overview of our results and compare them to outlier risk sharing. In Table 1, we show which type arises for each age-gender cell if we allow average cost to increase by 5 %. In addition, we show the percentage reduction in the difference of average costs for both optimal cost sharing and outlier risk sharing.

Figure 7 illustrates the effectiveness of the optimal cost sharing formula for a 1 to 10% increase in total cost. On average, the difference in expected costs can be decreased by 13.18 % if we allow costs to increase by 5% and by 23.12% for a 10% increase in costs. The maximum decrease is 30.73% and 67.87% respectively.

Figure 8 shows the same graph for outlier risk sharing. We find that this type of cost sharing is much less effective than our method. For example, if costs are

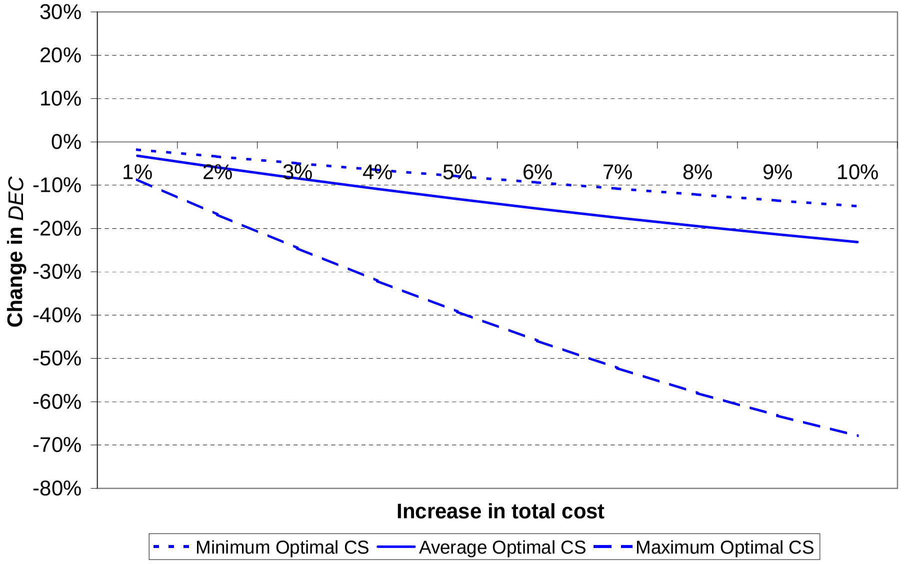

Figure 7: Reduction in the difference in expected costs with optimal cost sharing

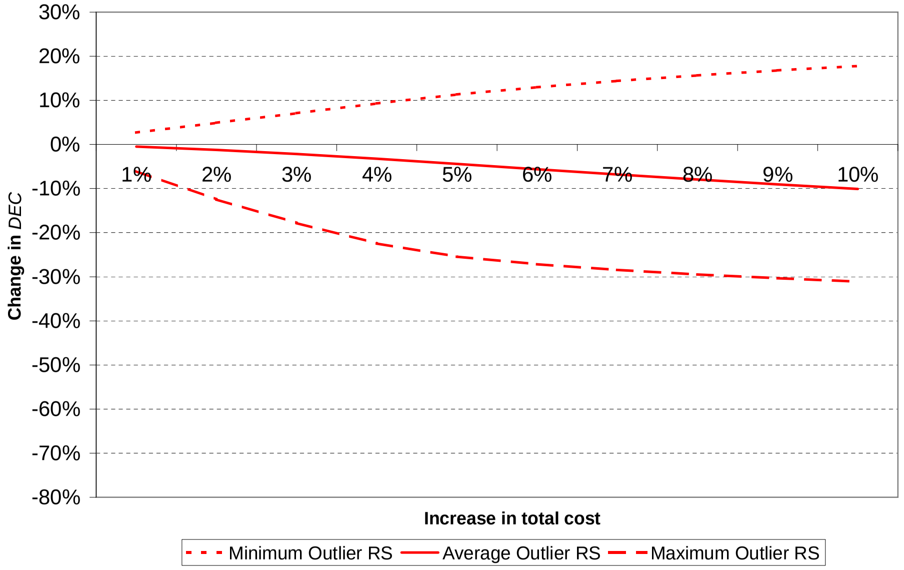

Figure 8: Reduction in the difference in expected costs with outlier risk sharing

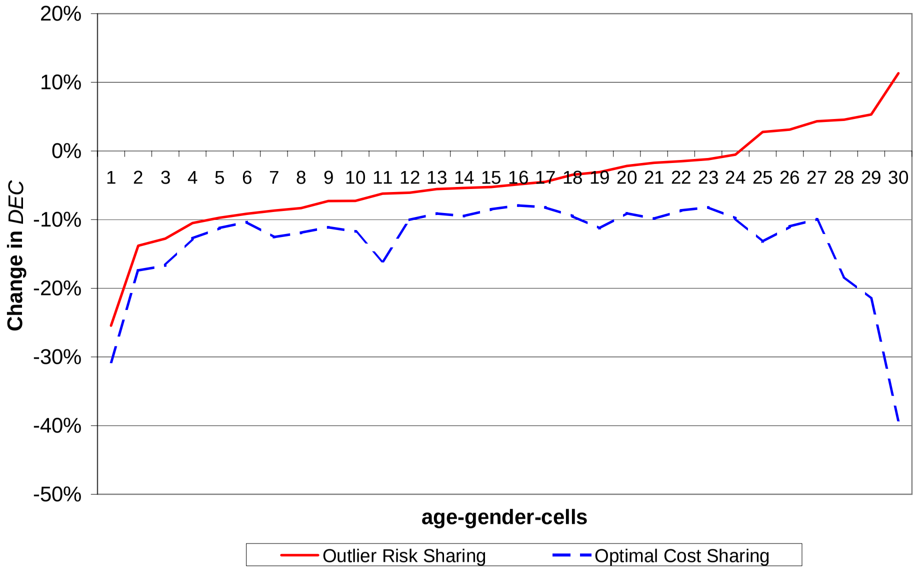

Figure 9: Optimal cost sharing and outlier risk sharing compared

allowed to increase by 5%, then the average reduction in the difference in expected costs is only 4.43% compared to 13.18% under optimal cost sharing. In addition, we find that the difference in costs increases under outlier risk sharing in 5 of 30 age-gender-cells. This shows that the scenario in our example is also possible in practice.

Finally, we show in Figure 9 the results for all 30 age-gender-cells for a 5%-increase in total costs for optimal cost sharing and outlier risk sharing. The cells were arranged by the magnitude of the reduction in the difference in expected costs of outlier risk sharing. We find that optimal cost sharing performs well when outlier risk sharing performs well. This is no surprise because optimal cost sharing can always mimic outlier risk sharing. However, it performs also well when outlier risk sharing performs very badly. The difference in expected cost can be 45% lower if optimal cost sharing is used instead of outlier risk sharing. In these cases, we find that cost sharing of type one is optimal where the threshold above which there is no additional cost sharing is low. The optimal cost sharing formula is therefore just the opposite of outlier risk sharing.

Cell	Age-Gender-Bracket	Number of observations	Percentage Hospitalized in 1998	Average Cost Hospitalized	Average Cost Non-Hosp.	Average Cost Ratio	*DEC*-Change in % Optimal Cost-sharing	*DEC*-Change in % Outlier Risk-sharing	Type
F1	Women 18-25	5902	6.2	4228	1392	3.04	-30.73	-25.43	3
F2	Women 26-30	4289	11.6	3570	1958	1.82	-16.01	-6.21	2
F3	Women 31-35	5372	14.9	3532	2103	1.68	-11.93	-8.33	2
F4	Women 36-40	5678	10.9	3934	1995	1.97	-10.05	-6.08	2
F5	Women 41-45	5262	8.7	5703	2060	2.77	-12.75	-10.50	2
F6	Women 46-50	5103	8.5	5671	2222	2.55	-16.66	-12.77	3
F7	Women 51-55	5049	8.7	6611	2582	2.56	-12.56	-8.69	3
F8	Women 56-60	4592	9.6	6805	2730	2.49	-11.23	-9.72	2
F9	Women 61-65	3907	11.5	7179	3200	2.24	-8.49	-5.25	2
F10	Women 66-70	3536	14.8	9707	3780	2.57	-11.10	-7.28	3
F11	Women 71-75	3224	18.1	10832	4526	2.39	-8.21	-4.50	1
F12	Women 76-80	2871	21.9	12675	5624	2.25	-9.80	-0.53	1
F13	Women 81-85	2022	27.5	15494	6476	2.39	-13.23	+2.77	1
F14	Women 86-90	1564	35.9	18501	8634	2.14	-18.31	+4.56	1
F15	Women 91+	803	48.9	22123	11733	1.89	-39.21	+11.32	1
M1	Men 18-25	5738	3.6	4881	854	5.72	-17.42	-13.80	3
M2	Men 26-30	3730	4.9	4466	1044	4.28	-9.06	-2.17	2
M3	Men 31-35	4609	4.3	9072	1154	7.86	-11.73	-7.26	3
M4	Men 36-40	4697	5.8	6381	1364	4.68	-7.94	-4.87	2
M5	Men 41-45	4363	5.6	5907	1482	3.98	-9.10	-5.55	2
M6	Men 46-50	4109	6.7	5764	1625	3.55	-9.50	-3.46	1
M7	Men 51-55	4240	7.8	7614	1928	3.95	-10.33	-9.15	3
M8	Men 56-60	3769	9.4	6774	2587	2.62	-9.87	-1.72	1
M9	Men 61-65	2920	10.8	9519	3203	2.97	-11.28	-3.06	2
M10	Men 66-70	2351	15.1	8680	4108	2.11	-9.96	+4.33	1
M11	Men 71-75	1852	19.1	11374	5405	2.10	-9.53	-5.39	3
M12	Men 76-80	1389	23.0	10754	5780	1.86	-8.22	-1.19	1
M13	Men 81-85	805	26.8	13599	7220	1.88	-11.00	+3.11	1
M14	Men 86-90	489	30.9	16984	8528	1.99	-8.66	-1.49	1
M15	Men 91+	185	37.3	19790	10907	1.81	-21.51	+5.32	1
	Average	3480	15.6	9271	3940	2.87	-13.18	-4.43	–

Table 1: Age-gender cells and results for a 5 % increase in costs

5 Conclusion

Cost sharing can be an important method to avoid risk selection in health insurance markets. In this paper, we derived the optimal cost sharing function based on a model in which insurers influence the cost of health care with their organizational activities. We found that on theoretical grounds there is no reason that the formula will be similar to outlier risk sharing which is advocated in the literature and applied in Germany. Our empirical analysis confirmed this result. We showed that the optimal cost sharing formula usually looks very different from outlier risk sharing. It was also much more effective. For a five percent increase in costs, we found that the mean decrease in the average cost difference between the two groups is almost three times larger if optimal cost sharing is used instead of outlier risk sharing. In addition, outlier risk sharing may actually increase the incentives for risk selection because it can increase the difference in average costs between the groups.

We think that our results call for a reconsideration of the current practice of cost sharing. Although we made a number of specific assumptions, we think that our method can be developed further to be applied in practice. The theoretical model could be extended to allow for multiple risk types and other types of cost functions. On the empirical side, our method requires a representative data set which contains information on the characteristics which insurers use to risk select and information on the cost function of health insurers. If this data is made available, cost sharing may become a powerful tool to prevent risk selection.

Appendix

Since we allow for $F_i(0) > 0$ it is not possible to set up the Hamiltonian and apply the maximum principle to solve our isoperimetric dynamic optimization problem with free starting and end points. In the following, we therefore solve the complete problem. To save on notation, we define $\tilde{H}(m) = \tilde{F}_h(m) - \tilde{F}_l(m)$ and $\tilde{h}(m) = \tilde{f}_h(m) - \tilde{f}_l(m)$. Then problem (P2) is equivalent to the maximization problem

$$\max_{r(.)} \; r(0)\tilde{H}(0) + \int_0^{\overline{C}} r(C)\tilde{h}(C)\,\mathrm{d}C \tag{A.1}$$

s.t.

$$r(0)\tilde{G}(0) + \int_0^{\overline{C}} r(C)\tilde{g}(C)\,\mathrm{d}C = 0 \tag{A.2}$$

$$\int_0^{\overline{C}} r'(C)C\tilde{g}(C)\,\mathrm{d}C = k(\tilde{e})c(\tilde{e}) \tag{A.3}$$

$$0 \le r'(C) \le 1 \tag{A.4}$$

$$r(0),\; r(\overline{C}) \;\text{ free} \tag{A.5}$$

Now replace constraint (A.2) by

$$K(C) = \int_0^{C} r(s)\tilde{g}(s)\,\mathrm{d}s \;\text{ with }\; K'(C) = r(C)\tilde{g}(C), K(0) = 0 \text{ and } K(\overline{C}) = -r(0)\tilde{G}(0).$$

Furthermore set $r'(C) = u(C)$ and replace (A.3) by

$$L(C) = \int_0^{C} u(s)s\tilde{g}(s)\,\mathrm{d}s \;\text{ with }\; L'(C) = u(C)C\tilde{g}(C),\; L(0) = 0 \text{ and } L(\overline{C}) = \tilde{\alpha}\tilde{c}.$$

Therefore the problem is

$$\max_{r(.)} \; r(0)\tilde{H}(0) + \int_0^{\overline{C}} r(C)\tilde{h}(C)\,\mathrm{d}C \tag{A.6}$$

subject to

$$\begin{aligned}
&K'(C) = r(C)\tilde{g}(C), && K(0) = 0, && K(\overline{C}) = -r(0)\tilde{G}(0)\\
&L'(C) = u(C)C\tilde{g}(C), && L(0) = 0, && L(\overline{C}) = k(\tilde{e})c(\tilde{e})\\
&K'(C) = r(C)\tilde{g}(C), && K(0) = 0, && K(\overline{C}) = -r(0)\tilde{G}(0)\\
&r(0), r(\overline{C}) \;\text{ free}
\end{aligned}$$

We can now set up the Lagrangian

$$\begin{aligned} L \;=\;& r(0)\tilde{H}(0)+\int_{0}^{\overline{C}}\Big\{r(C)\tilde{h}(C)+\lambda(C)[u(C)-r'(C)] \\ &+\mu(C)[r(C)\tilde{g}(C)-K'(C)]+\eta(C)[r'(C)C\tilde{g}(C)-L'(C)]\Big\}\mathrm{d}C \\ &+\gamma_1K(0)+\gamma_2[r(0)\tilde{G}(0)+K(\overline{C})]+\gamma_3L(0)+\gamma_4[k(\tilde{e})c(\tilde{e})-L(\overline{C})]. \end{aligned} \tag{A.7}$$

Note that $\eta(C)$ is the Lagrange multiplier associated with the incentive constraint (A.3). Integrating $\lambda(C)r'(C)$, $\mu(C)K'(C)$ and $\eta(C)L'(C)$ by parts we obtain

$$\begin{aligned} L \;=\;& r(0)\tilde{H}(0)+\int_{0}^{\overline{C}}\Big\{r(C)\tilde{h}(C)+\lambda(C)u(C)+\lambda'(C)r(C) \\ &+\mu(C)r(C)\tilde{g}(C)+\mu'(C)K(C)+\eta(C)r'(C)C\tilde{g}(C)+\eta'(C)L(C)\Big\}\mathrm{d}C \\ &-[\lambda(\overline{C})r(\overline{C})-\lambda(0)r(0)] \\ &-[\mu(\overline{C})K(\overline{C})-\mu(0)K(0)]-[\eta(\overline{C})L(\overline{C})-\eta(0)L(0)] \\ &+\gamma_1K(0)+\gamma_2[r(0)\tilde{G}(0)+K(\overline{C})]+\gamma_3L(0)+\gamma_4[k(\tilde{e})c(\tilde{e})-L(\overline{C})]. \end{aligned} \tag{A.8}$$

The first differential is

$$\begin{aligned} \Delta L \;=\;& \int_{0}^{\overline{C}}\Big\{[\tilde{h}(C)+\lambda'(C)+\mu(C)\tilde{g}(C)]\Delta r(C)+[\lambda(C)+\eta(C)C\tilde{g}(C)]\Delta u(C) \\ &+\mu'(C)\Delta K(C)+\eta'(C)\Delta L(C)\Big\}\mathrm{d}C \\ &+[\tilde{H}(0)+\lambda(0)+\gamma_2\tilde{G}(0)]\Delta r(0)-\lambda(\overline{C})\Delta r(\overline{C})+[-\mu(\overline{C})+\gamma_2]\Delta K(\overline{C}) \\ &+K(0)\Delta\gamma_1+[r(0)\tilde{G}(0)+K(\overline{C})]\Delta\gamma_2+L(0)\Delta\gamma_3+[k(\tilde{e})c(\tilde{e})-L(\overline{C})]\Delta\gamma_4. \end{aligned} \tag{A.9}$$

This yields the following conditions for optimality

$$\tilde{h}(C)+\lambda'(C)+\mu(C)\tilde{g}(C)=0 \tag{A.10}$$

$$\lambda(C)+\eta(C)C\tilde{g}(C)\begin{cases} >0 & \Rightarrow \quad u(C)=1 \\ =0 & \Rightarrow \quad 0\le u(C)\le 1 \\ <0 & \Rightarrow \quad u(C)=0 \end{cases} \tag{A.11}$$

$$\mu'(C)=0 \quad \text{which implies} \quad \mu(C)=\overline{\mu} \tag{A.12}$$

$$\eta'(C)=0 \quad \text{which implies} \quad \eta(C)=\overline{\eta} \tag{A.13}$$

$$\lambda(0)=-\tilde{H}(0)-\gamma_2\tilde{G}(0) \tag{A.14}$$

$$\lambda(\overline{C})=0 \tag{A.15}$$

$$\gamma_2 = \mu(\overline{C}) \quad \text{which implies} \quad \gamma_2 = \overline{\mu}. \tag{A.16}$$

$$K(0) = 0 \tag{A.17}$$

$$r(0)\tilde{G}(0) + K(\overline{C}) = 0 \tag{A.18}$$

$$L(0) = 0 \tag{A.19}$$

$$L(\overline{C}) = k(\tilde{e})c(\tilde{e}). \tag{A.20}$$

Integrating (A.10) yields

$$\begin{aligned}\lambda(C) &= \lambda(0) + \int_0^C \lambda'(s)\,\mathrm{d}s \\ &= -\tilde{H}(0) - \overline{\mu}\tilde{G}(0) + \int_0^C -\tilde{h}(s) - \overline{\mu}\tilde{g}(s)\,\mathrm{d}s \\ &= -\tilde{H}(C) - \overline{\mu}\tilde{G}(C). \end{aligned} \tag{A.21}$$

With

$$0 = \lambda(\overline{C}) = -\tilde{H}(\overline{C}) - \overline{\mu}\tilde{G}(\overline{C}) = 0$$

we get $\overline{\mu} = 0$ which simplifies (A.21) to

$$\lambda(C) = -\tilde{H}(C).$$

Inserting into (A.11) we obtain

$$-\tilde{H}(C) + \eta(C)C\tilde{g}(C) \begin{cases} > 0 & \Rightarrow \quad u(C) = r'(\overline{C}) = 1 \\ = 0 & \Rightarrow \quad 0 \leq u(C) = r'(\overline{C}) \leq 1 \\ < 0 & \Rightarrow \quad u(C) = r'(\overline{C}) = 0 \end{cases}. \tag{A.22}$$

which is equivalent to condition (13). Now $\overline{\eta}$ needs to be chosen such that (A.20) is satisfied. This guarantees

$$\int_0^{\overline{C}} r'(C)C\tilde{g}(C)\,\mathrm{d}C = k(\tilde{e})c(\tilde{e}).$$

Finally $r(0)$ is set such that (A.18) is satisfied which implies

$$r(0)\tilde{G}(0) + \int_0^{\overline{C}} r(C)\tilde{g}(C)\,\mathrm{d}C = 0. \tag{A.23}$$

References

Arrow, K. (1974). “Optimal Insurance and Generalized Deductible,” *Scandinavian Actuarial Journal* 57, 1–42.

Chiang, A. (1992). *Elements of Dynamic Optimization.* McGraw-Hill: New York.

Frank, R., J. Glazer, and T. McGuire (2000). “Measuring adverse selection in managed health care,” *Journal of Health Economics* 19, 829–854.

Glazer, J., and T. McGuire (2000). “Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care,” *American Economic Review* 90, 1055–1071.

Glazer, J., and T. McGuire (2002). “Setting Health Plan Premiums to Ensure Efficient Quality in Health Care: Minimum Variance Optimal Risk Adjustment,” *Journal of Public Economics* 84, 153–173.

Kamien, M., and N. Schwartz (1991). *Dynamic Optimization: the Calculus of Variations and Optimal Control in Economics and Management.* No. 31 in Advanced Textbooks in Economics North-Holland: Amsterdam, 2 edn.

Marchand, M., M. Sato, and E. Schokkaert (2003). “Prior Health Expenditures and Risk Sharing with Insurers Competing on Quality,” forthcoming RAND Journal of Economics.

Newhouse, J. (1994). “Patients at Risk: Health Reform and Risk Adjustment,” *Health Affairs* 13, 132–146.

Pauly, M. (1984). “Is Cream-Skimming a Problem for the Competitive Medical Market?,” *Journal of Health Economics* 3, 87–95.

Raviv, A. (1979). “The Design of an Optimal Insurance Policy,” *American Economic Review* 69, 84–96.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman & Hall: London.

Terrell, G., and D. Scott (1992). “Variable Kernel Density Estimation,” *Annals of Statistics* 20, 1236–1265.

van Barneveld, E., L. Lamers, R. van Vliet, and W. van de Ven (2001). “Risk sharing as a supplement to imperfect capitation: a tradeoff between selection and efficiency,” *Journal of Health Economics* 20, 147–168.

van de Ven, W., and R. Ellis (2000). “Risk Adjustment in Competitive Health Plan Markets,” in A.J. Culyer, and J.P. Newhouse (ed.), *Handbook of Health Economics*, vol. 1a, pp. 755–845 Elsevier: Amsterdam.

van de Ven, W., and R. van Vliet (1992). “How Can We Prevent Crem Skimming in a Competitive Health Insurance Market?,” in P. Zweifel, and H.E. Frech III (ed.), *Health Economics Worldwide* pp. 23–46 Kluwer: Dordrecht.