Medical Savings Contract or Deductible Insurance?

Ronny Klein*

February 2003

Abstract

A medical savings contract (MSC) is a health insurance which includes a savings account: when ill, the insured pays the medical costs out of the account until it is depleted. All further costs are taken by the insurer. This paper compares the MSC and a common deductible insurance (DI) using a two-period-model. Firstly, we derive conditions under which an expected utility maximizer will prefer a MSC over a DI. Secondly, we show that risk aversion could lead the insured to take less long-term disease prevention activities under a MSC than under a DI. In addition, we discuss the possibility of a private savings plan.

JEL classification: D80, D91, G22, I10
Keywords: Insurance, Medical Savings Account, Deductible, Moral Hazard

1 Introduction

Individual savings accounts as a supplement to a social security insurance are widely accepted within the context of pension funding systems as a way to temper the financial burden of the aging society. The introduction of the so called *Riester Rente* in Germany in the year 2002 is one of the recent examples for the popularity of this idea. In 2002, Chile introduced individual unemployment accounts which are complemented by a public fund in the case of depletion.[1] But, also in health care financing, there is growing interest in introducing individual health accounts in order to create incentives to utilize health care more prudently and to encourage disease prevention activities.

In 1996 the 104th Congress released the *Health Insurance Portability and Accountability Act*[2] and established a pilot program for medical savings

*Martin-Luther-University Halle-Wittenberg, Department of Economics, 06099 Halle, Phone: +49/345/5523443, Fax: +49/345/5527127, Email: r.klein@wiwi.uni-halle.de.

[1] See Feldstein and Altman (1998) and Orszag and Snower (2002) for an analysis of unemployment accounts.

[2] HIPAA, Public Law 104-191, August 21, 1996.

accounts (MSA) in the USA. The Internal Revenue Service (2001) defines a MSA as follows: "An Archer MSA is a tax-exempt trust or custodial account with a financial institution (like a bank or an insurance company) in which you can save money for future medical expenses. This account must be used in conjunction with an HDHP."[3] In particular, the tax exemption of contributions to a MSA should be an incentive to sign such a contract. Bunce (2001) gives an overview of the development of MSAs in the USA. Although during the pilot program the public interest was surprisingly low, according to the concluding report of the GAO (1998), the program has been extended until the end of 2002.

In Singapore, the so called *Medisave accounts* were already established nationwide in 1984. Massaro and Wong (1995, 1996) argue that this is the reason for the relative low health expenditure of 3.1% of the GNP in comparison with other Asian countries and all the more with western industrial nations.

According to Matisonn (2000), MSAs became a popular health insurance in South Africa very soon after the deregulation of the insurance market in 1994. While the market share of HMO-like contracts is about 4%, MSA-based contracts are rather widespread with 51%.

Yip and Hsiao (1997) describe the experiences of the Chinese with a MSA-based social insurance experiment in two cities, and concludes that there is only little empirical evidence on the impact of MSAs on overall health care expenditures.

In the USA advantages and disadvantages of MSAs are discussed at political and academic level. In Europe, however, the debate does not exist until now.[4] Furthermore, there is only little theoretical insight in the decision-making process of an individual when setting up a MSA. Heffley and Miceli (1998) analyze MSAs only as an example within their framework of incentive-based health care plans. Klein (2002) extends the standard economic model of a deductible insurance to a two-period model with a savings account, and characterizes the optimal deposit.

In this paper, we will analyze the behavior of an individual when facing a decision between a standard deductible insurance and a medical savings contract.[5] While using the two-period model of Klein (2002), we want to answer the following questions:

[3]HDHP: high deductible health plan. High deductible means higher than with typical health plans.

[4]Actually, there are only some comments as far as I known. For the USA one can mention: American Academy of Actuaries (1995a, 1995b), Bunce (2001), Friedman (2001), Moon et al. (1996), Ozanne (1996), Pauly and Herring (2000), Pauly and Goodman (1995), Scandlen (1998), Stano (1981), and Zabinski et al. (1999), but the list does not claim to be complete.

[5]MSC means the combination of a MSA and an insurance. With MSA, we only denote the account itself.

- When will an expected utility maximizer prefer a medical savings contract over a deductible insurance?
- Could a medical savings contract also be replaced by a private savings plan?
- Under which policy we can expect more long-term disease prevention activities?

The paper is organized as follows: in the second section, we introduce the general model of a MSC. In the third section, we determine the conditions which must be met in order to make a MSC superior to a DI. Furthermore, we compare a MSC with a private savings plan. In the forth section, we introduce the possibility for the insured to take long-term disease prevention activities and compare the levels under a MSC and a DI, respectively. The fifth section concludes the paper.

2 The Model

Two types of individual health accounts based insurances can be distinguished: the US-type and the Singapore-type MSC. The main difference lies in the flexibility of the overall deductible: while this one is fixed for US-type MSC, the maximum of one's own contribution to the health care costs under a Singapore-type MSC depends on the balance of the MSA. The above mentioned Archer MSA and the MSAs in South Africa belong to the US-type MSCs. Medisave in conjunction with Medishield and Medifund in Singapore can be interpreted as a MSC with a flexible overall deductible: the higher the savings in the Medisave account the higher can be the insured's copayments. Only, when the Medisave account is depleted – because of the 20% copayments of the Medishield insurance– and the insured is not able to cover the remaining costs of treatment out-of-pocket, Medifund will step into the breach.[6] The Chinese MSA and the individual health accounts proposed by Stano (1981) belong also to the Singapore-type MSC. Both approaches consist of individual MSAs and a social insurance which covers the health care costs in the case of a depleted MSA.

Differences in the flexibility of the overall deductible lead directly to differences in the regularity of the insured's payment into her MSA. If the overall deductible is fixed, there is no need to require regular deposits. It can be left to the insured herself to decide on. However, if the overall deductible depends on the balance of the account, the premium of the additional insurance will depend on it, too. Therefore, in order to calculate an expected

[6]This is, admittedly, a highly compressed representation of Singapore's health care financing system. For a detailed description, we want to refer to Massaro and Wong (1996) and the Internet site of Singapore's Ministry of Health: http://www.gov.sg/moh.

balance of the MSA, one need regular deposits into the account. This can be observed in Singapore and China where the insureds must pay a certain share of their monthly income into the MSA.

What we want to model here is the Singapore-type MSA. The reason is that a US-type MSA is not very much different from a standard deductible insurance due to the fixed overall deductible. In a previous paper (Klein 2002), we outlined such a model and we will use this approach here again.

For simplicity, we use a two-period model.[7] These periods can be either interpreted as policy-years or as phases of one's life like youth and age. At first, we define the risk to be insured against.

Assumption A. 1 (The Risk). *Let Ω be the set of all possible diseases.*

i) $A : \Omega \to \mathbb{R}^+$ *is a continuous random variable,*

ii) $F(a_1, a_2)$, *with* $a_1, a_2 \in A$, *is a joint distribution function, where the index indicates the period.*

iii) F *is common knowledge and insurable.*

The elements of A can be interpreted as the whole follow-up costs of a disease. Thus, F is the joint distribution function of the insured's medical expenses in period one and two. With *insurable*, we mean that, firstly, there is a real risk involved, i.e., F is not degenerated, and, secondly, an insurance policy against the losses born by F would not be too expensive, and a potential insurance buyer can afford the policy.

Assumption A. 2 (Individual's Preferences). *Let u be the insured's utility function, and $w \in \mathbb{R}^+$ the insured's wealth with $u : w \to \mathbb{R}$. Furthermore, denote the discount factor with $\rho \in (0, 1]$.*

i) $u \in \Upsilon := \{u : \; u' > 0, u'' < 0 \quad \forall w \in \mathbb{R}^+\}$,

ii) $v := \mathbb{E}(u_1 + \rho u_2)$, *where the index only indicates the period, without altering the utility function itself.*

Firstly, we assume a strict risk averse individual whose welfare per period can be represented with a continuous, strict concave and two times differentiable utility function. In addition, the insured's utility is supposed to depend on her wealth, only. So, for simplicity, we ignore the possible impact of an illness on the insured's welfare. Secondly, we assume that the preferences over different states of the world, i.e., as a consequence of different insurance policy parameters, can be represented by the sum of expected utilities per period.

[7] Similar single-period models of a deductible insurance can be found in Eeckhoudt et al. (1991), Gould (1969), Mossin (1968), and Schlesinger (1981).

The supply side is, for the sake of simplicity, only rudimentary modeled: An insurance company offers the insurance buyer a policy which yields zero expected profits; either due to competition or because of public regulations. Thus, the insurance buyer can be insured according to a premium function which relates the expected indemnity payments and costs of the insurance company to the policy premium.

Assumption A. 3 (Premium Function). *Let p be the premium per period and i the indemnity payments born by the policy. In addition, denote the insurance loading with $\lambda \geq 0$ and the interest factor with $r \geq 1$. Then, the premium function is*

$$p = \frac{1+\lambda}{1+r}\mathbb{E}(ri_1 + i_2), \tag{1}$$

with p constant in period one and two.

In a single-period model, the premium function would be $p = (1+\lambda)\mathbb{E}i$. The denominator in (1), therefore, equalizes the per period premiums. In addition, one has to include the interest costs or gains, respectively, of the indemnity payments in period one. Hence, these have to be multiplied by the interest factor.

The Policies With a common deductible insurance, we mean an insurance where, in each period, the insured pays the medical expenses up to the amount of the deductible. All costs above are covered by the insurer. The deductible is fixed and from equal size in each period.

A MSC is a combination of a MSA and an insurance. The insured agrees to pay a certain amount of money into the MSA in each period. Withdrawals from the MSA are only allowed for medical purposes. Medical expenses are firstly paid out of the MSA. If the savings are not enough, the insurer pays the rest. After the policy runs out, the remaining savings are distributed to the insured.

It is rather cumbersome to compare two different policies directly. Thats why, we create an artificial combined policy with two parameters: the maximum out-of-pocket payment per period (MOP) $d \geq 0$ and the rate of saving $\delta \in [0, 1]$ of unused MOP which will be saved in the MSA. So, a chosen $\delta = 0$ means the policy is a common deductible insurance. And $\delta = 1$ means the insurance buyer chooses a pure MSC. However, a rate of saving between zero and one is also possible and can also be interpreted as a kind of MSC. In each policy year the timetable should be as follows:

1. The medical expenses of the period become known and are withdrawn from the MSA.

2. If the savings are not enough, the insured has to pay out-of-pocket.

3. If the costs exceed the MOP, too, the insurer pays the remaining costs.

4. If the medical expenses are lower than the MOP, the insured pays the share of unused MOP into the MSA according to the rate of saving.

5. In the last policy year, the remaining savings will be paid back to the insured.

Table 1 shows how such a combined policy will work under different rates of saving. One can easily verify that both of the above described policies can be represented by the combined policy.

1.	Rate of Saving	$\delta = 0$		$\delta = 0.5$		$\delta = 1$	
2.	Policy Year	1.	2.	1.	2.	1.	2.
3.	MSA Balance, 1. Jan	0	0	0	400	0	800
4.	Medical Expenses	200	1500	200	1500	200	1500
5.[a]	Indemnity payments	0	500	0	100	0	0
6.[b]	MSA Balance, 31. Dec	0	0	400	0	800	300

[a] $5. = \max\{4. - 3. - d, 0\}$.
[b] If $4. \geq 3.$, $6. = \max\{\delta(3. + d - 4.), 0\}$. If $4. < 3.$, $6. = 3. - 4. + \delta d$.

Table 1: Examples of the combined policy with $d = 1000$ and different rates of saving.

Denote with l (low) the case when the insurer's indemnity payments in the period are zero due to relative low medical expenses, and with h the opposite. Therefore, in a two-period model, we can distinguish four different disease histories: ll, lh, hl and hh.

In Figure 1, each point reflects a pair of medical costs in period one and two. In the case of ll, the insured has to bear all of the expenses. This area increases with δ.

To abstract from other savings strategies, we impose two additional assumptions:

Assumption A. 4. *Let y be the insured's income per period. The insured's income is constant over time, i.e., $y_1 = y_2 \equiv y$.*

Assumption A. 5. $\rho r = 1$.

Now, we can calculate the expected utility in the case of a certain disease history:

$$\begin{aligned} v_{ll} &= \mathbb{E}_{ll}[u(y - p - a_1 - \delta[d - a_1]) + \rho u(y - p - a_2 + r\delta[d - a_1])], \\ v_{hl} &= \mathbb{E}_{hl}[u(y - p - d) + \rho u(y - p - a_2)], \\ v_{lh} &= \mathbb{E}_{lh}[u(y - p - a_1 - \delta[d - a_1]) + \rho u(y - p - d)], \\ v_{hh} &= \mathbb{E}_{hh}[u(y - p - d) + \rho u(y - p - d)]. \end{aligned} \tag{2}$$

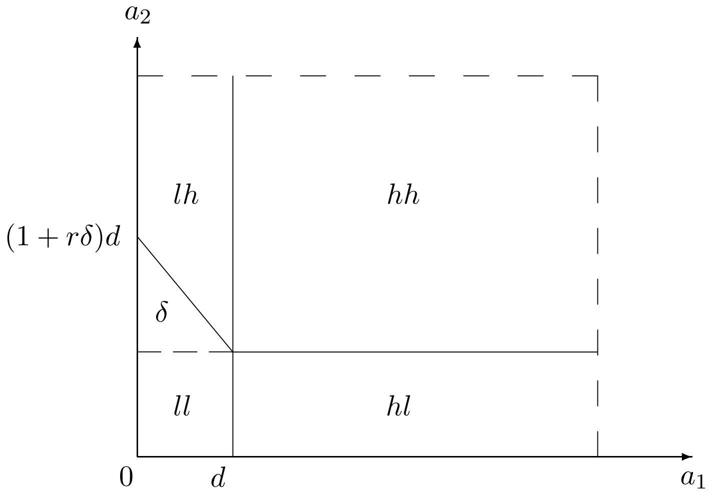

Figure 1: Possible disease histories in a two-period model.

The insured's overall expected utility is, therefore, the sum of the utility of all possible states of the world

$$v(d,\delta) = v_{ll} + v_{lh} + v_{hl} + v_{hh}. \tag{3}$$

At last, we assume that the insurance loading – and, therefore, the costs of the insurer – is positive and independent of the rate of saving.

Assumption A. 6. $\lambda > 0$ *constant* $\forall \delta$.

Thus, the premium wont be actuarial fair. This gives an incentive to take no full coverage. Furthermore, the discrimination between MSC and DI is not supposed to be based on different cost structures.

3 The Expected Utility Criterion

3.1 MSC vs. Deductible Insurance

In this section, we want to compare MSC and DI from the insured's point of view. Can there be any reason found why a potential insurance buyer will choose a MSC over a DI? The problem for an expected-utility maximizer when facing the choice of a combined policy is as follows:

$$\begin{aligned} &\max_{d,\delta} v(d,\delta) = \mathbb{E}(u_1 + \rho u_2) \\ &\text{subject to } p(d,\delta) = \frac{1+\lambda}{1+r}\mathbb{E}(ri_1 + i_2). \end{aligned} \tag{4}$$

Now, we can answer the question whether a MSC is superior to a DI or vice versa: if the insurance buyer chooses $\delta^* = 0$, she prefers the pure deductible insurance, where the * denotes the solution to problem (4). If she demands a $\delta^* > 0$, she prefers a MSC. However, you have to know the utility function and the distribution of medical costs to predict the optimal δ directly. Instead of specifying this, we will follow the question for what types of insurance buyers you can expect a choice of $\delta^* = 0$ and for whom not.

Let us firstly show that an insurance buyer will always choose a positive maximum out-of-pocket payment per period whenever the insurance premium is not actuarial fair.

Lemma 1 (Positive MOP). *Suppose A.1-A.6; then* $d^* > 0$*.*

Proof. The lemma is proved if the first derivative of $v(d, \delta)$ w.r.t. d at the point $d = 0$ is positive for every $\lambda > 0$. Note that with a zero MOP only hh is possible and $u_1' = u_2' = u'(y - p)$ in all situations.

$$v_d'|_{d=0} = -p_d'\mathbb{E}(u_1' + \rho u_2')] - \mathbb{E}_{hh}(u_1' + \rho u_2'), \tag{5}$$

with

$$-p_d' = \frac{1+\lambda}{1+r}[\mathbb{E}_{hh}(1+r)]. \tag{6}$$

Let $v' := \mathbb{E}(u_1' + \rho u_2')$. Setting (6) into (5) yields $v_d'|_{d=0} = \lambda v' > 0$ if and only if $\lambda > 0$. Therefore, the expected utility at $d = 0$ cannot be maximal. □

Now, let $\mathcal{D}_0 := \{d : v_d' \geq 0 \;\&\; \delta = 0\}$ and $\mathcal{D}_0^* := \{d : v \text{ maximal} \;\&\; \delta = 0\}$. Then, clearly $\mathcal{D}_0^* \subset \mathcal{D}_0$. Below, we will need the following result:

Lemma 2. *Suppose A.1-A.6 and let* $u_d' = u'(y - p - d)$*; then, for every* $d \in \mathcal{D}_0$*:* $r\frac{1+\lambda}{1+r}v' \geq u_d'$*.*

Proof. Setting $v_d'|_{\delta=0} \geq 0$ and multiplying by r yields the result. See appendix A.1 for the details. □

With that in mind, we can formulate a condition such that for every positive $d \in \mathcal{D}_0$ the insured can increase her welfare when choosing a positive rate of saving.

Lemma 3 (Positive Rate of Saving). *Suppose A.1-A.6; then, for every given positive deductible* $d \in \mathcal{D}_0$*, a sufficient condition for* $\delta^* > 0$ *is*

$$\mathbb{E}_{ll}[(d - a_1)(u_1' - r\rho u_2')] \leq \mathbb{E}_{lh}[(d - a_1)(u_d' - u'(y - p - a_1)] \tag{7}$$

at the point $\delta = 0$*.*

Proof. In the same manner as before, we are done if we can show that $v'_\delta > 0$ at the point $\delta = 0$ when (7) applies.

$$v'_\delta|_{\delta=0} = -\, p'_\delta v' - \mathbb{E}_{ll}[(d-a_1)(u'_1 - r\rho u'_2)] - \mathbb{E}_{lh}[(d-a_1)u'_1], \tag{8}$$

with

$$-p'_\delta = \frac{1+\lambda}{1+r}\mathbb{E}_{lh}[r(d-a_1)]. \tag{9}$$

Setting (9) into (8) yields

$$\begin{aligned} v'_\delta|_{\delta=0} = &-\mathbb{E}_{ll}[(d-a_1)(u'_1 - r\rho u'_2)] \\ &+ \mathbb{E}_{lh}\left\{(d-a_1)\left[r\frac{1+\lambda}{1+r}v' - u'_1\right]\right\}. \end{aligned} \tag{10}$$

Because of Lemma 2 and $r\rho = 1$, equation (10) can be transformed to

$$\begin{aligned} v'_\delta|_{\delta=0} \geq &-\mathbb{E}_{ll}[(d-a_1)(u'_1 - u'_2)] \\ &+ \mathbb{E}_{lh}\left\{(d-a_1)\left[u'_d - u'(y-p-a_1)\right]\right\}. \end{aligned} \tag{11}$$

Due to risk aversion $v'_\delta > 0$ at the point $\delta = 0$ if (7) applies, and the insured can increase her expected utility when choosing a positive rate of saving. □

This result leads us directly to our first main statement.

Proposition 1. *Suppose A.1-A.6; then, if (7) applies, a risk averse individual will always prefer a MSC over a pure deductible insurance, i.e.,* $d^* > 0$ *and* $\delta^* > 0$.

Proof. This result follows immediately from the three lemmas above: Due to Lemma 1 $d^* > 0$; together with (7) and the fact that $d^* \in \mathcal{D}_0^* \subset \mathcal{D}_0$ all of the conditions for $\delta^* > 0$ are met. □

The reason for this result is the relative high income in period one when choosing a pure deductible solution and lh occurs. With a positive rate of saving, the insured can transfer some income from lh to other states of nature, i.e., where she has to bear more medical costs. There is no way to do that with a change of the deductible itself. An increase of δ, however, decreases the running income for consumption in the first period, immediately. This leads to a lower premium, which increases the insured's utility in all situations.

In the case of ll a rate of saving higher than zero leads only to wealth shifting from the first to the second period. This fact is represented by the lhs of (7). If the medical costs are independent and identical distributed, there is no difference between the marginal utilities in period one and two. That's why the second term in (7) becomes negative, how the following corollary states:

Corollary 1.1. *Suppose A.1-A.6 and independent and identical distributed medical costs in period one and two; then, a risk averse individual will ever prefer a MSC over a pure DI, formally*

$$A.1 - A.6,\ f(a_1, a_2) \equiv f(a_1)f(a_2) \quad \& \quad f(a_1) = f(a_2) \quad \Rightarrow \quad \delta^* > 0 \quad (12)$$

Proof. See appendix A.2. □

Only, if there is a higher weight on those cases where the medical costs in the first period are higher than in the second period equation (7) could not be met. In this case, the marginal costs of an increase of δ in the first period could be higher than the marginal gain in period two. If this effect is high enough, the above mentioned utility gain through consumption smoothing could be overwhelmed.

Until now, we only have seen that the rate of saving can be positive under certain circumstances. But, we do not know if a rate of saving of 50% or even 100% is optimal. One way to show that the latter will be preferred to all other rates of saving is to ensure that the expected utility is ever increasing in δ.

Proposition 2. *Suppose A.1-A.6; then, if the probability of lh is sufficiently higher than the probability of ll, a risk avers individual will ever prefer a pure MSC with $\delta^* = 1$.*

Proof. If we show that $v'_\delta > 0$ for all $\delta \in [0, 1]$ and possible candidates for d^*, we are done because then δ^* will be increased until its upper boundary. This can be done by using a little altered version of Lemma 2 and 3. The details can be found in the appendix A.3. □

The relation between the probabilities of the illness histories ll and lh is crucial for the decision between a classic deductible insurance and a MSC. If lh, a MSC smooths the consumption stream better. However, in the case of ll, a MSC leads to relative low utility in the first period and a higher one in the second period due to the saving and reimbursement of the deductible. In contrast, a classic deductible insurance requires only to pay the medical costs in both periods, and the utilities do not differ as much. When hl or hh, the two policies do not differ in their associated consumption streams, thus, the probability of hl and hh should not matter.

A MSC allows to accumulate capital in healthy years which can be used later on – in years of sickness – to keep the premium affordable. If we interpret the periods as the *youth* and the *age*, the model predicts a high rate of saving due to the fact that the medical costs raise in the second part of one's lifetime for most of us, and, therefore, lh is more likely.

Arrow's Theorem The use of Arrow's (1963) theorem allows another way to think about the superiority of a policy. The theorem states that a policy cannot be optimal if the insured gets indemnity payments in a state of nature where the insured's wealth in another state of nature will be lower. With a DI, this would be exactly the case when lh and hl because the insured's wealth is in hh lower than in all the other states. Thus, in a two-period model a DI cannot be optimal. On the other side, with a pure MSC the insured's wealth in lh and hh is equal, and, therefore, the welfare should be higher. However, a pure MSC is also not optimal because the wealth in hl is higher than in lh and hh. The only optimal two-period policy would be a fixed overall deductible over the sum of the medical expenses in the two periods. This can be done, for instance, if the insured can go in debt on her MSA. However, how Cohen (2001) shows, such an aggregate deductible can lead to an increased moral hazard problem when the deductible is reached in the first period, and may, therefore, not be preferable.

3.2 MSC vs. Private Savings Plan (PS)

So, why do we need a new type of health care insurance? If savings in periods of good health can increase the insured's welfare, she can do it on her own: If healthy in the first period, save the unused deductible privately and increase the deductible in the second period by the savings. We will argue that this is inferior to a MSC because of the premium calculation. When setting up a MSC, the insured binds herself to save the unused deductible for future medical expenditures, so that the insurer *knows* about it. Thus, the insurer can smooth the premium. This is not possible with a private savings plan because, per definition, the insurer does not know about it. The insured has to face different premiums in the two periods which yields a lower expected utility due to risk aversion.

Proposition 3. *Suppose A.1-A.6 and $\delta = 1$; then, a risk averse individual will prefer a MSC over a PS.*

Proof. Note that the expected indemnity payments have to be the same for both plans because the insured imitates exactly the MSC when choosing the PS. Therefore, the expected aggregate premium payments and, thus, the expected aggregate wealth will also be equal.

Let p_{MSC} be the constant premium when setting up a MSC. Then in the first period under a PS, we have $p > p_{MSC}$ and in the second: $p \leq p_{MSC}$ due to the increased deductible. However, it is easy to see that with $\delta = 1$, the wealth in period one is never higher than in period two under both policies. Thus, the different premium payments under a PS increase only the difference between the first and second period wealth. But, if the expected aggregate wealth is equal, a risk averse individual will suffer a reduction in expected utility when the difference in wealth rises. Thus, a MSC will be preferred over a PS. □

Unfortunately, the same cannot be stated for $\delta < 1$ because the utility in period one can be higher than in two. Therefore, a reduction of u_1 and an increase of u_2 due to the differences in premiums when setting up a private savings plan can be superior to a fix premium. This is actually a different way to reduce the utility differences between the illness history lh and the others: one can increase δ within the frame of a MSC or save the unused deductible, privately.

4 Long-term Disease Prevention

Until now, we have only considered the case when the probability of an illness cannot be influenced. This assumption will now be dropped. Advocates of MSCs argue that such a policy will lead to greater responsibility by the insureds for their own health care costs. In this section, we will answer the question whether this can be expected or not.

We want to focus on ex ante moral hazard. It is rather known that a deductible insurance is not the best choice when facing ex post moral hazard due to the lack of incentives when the costs exceed the deductible.[8] This is also true for a MSC because after an illness has occurred there is no need to limit the costs above the MSA balance. The only difference lies in the size of the insured's own payments and the timing of those.

But, we argue that it is easy to establish a sort of co-insurance for a MSC and a classic deductible insurance. The insurer can interpret the overall deductible as a per capita maximum payment (CMP). Under this limit, there are every contract form possible, e.g., co-insurance. For instance, with a 90% coverage rate and a CMP of 1000 €, the insured has to pay until the medical costs exceed 10.000 €. So, with different coverage rates and CPMs, one ensures an incentive to keep the costs low for a wide range of diseases.

Another possibility to reduce the severity of the ex post moral hazard problem is to establish fixed amounts for certain kinds of diseases which will be paid out of the MSA. If the insured wishes a more expensive treatment, she has to pay it out-of-pocket.

Both of the above mentioned can be found in Singapore and are considered as very successful.[9]

Ex ante moral hazard In the two-period framework of our previous analysis, there are three different prevention technologies possible: actions

[8]See, e.g., Breyer and Zweifel (1997), Drèze (2002), and Winter (1992).

[9]See, e.g., Massaro and Wong (1995). Barr (2001) even argues that these elements are the main reason for Singapore's low health care expenditures.

which reduce the probability of a disease in the present period as well as actions which reduce the probability of an illness in the future period.[10] Let e_1, e_2 be the correspondent effort of the former ones and e for the latter. In addition, let π_t be the probability of a disease in period t ($\pi_t = Pr\{a_t > 0\}$). The following assumption will be made:

Assumption A. 7 (Prevention Activities). *Prevention activities which reduce the probability of diseases in the future are much more powerful than activities which reduce the present disease probability, i.e.,*

$$\pi_t'(e_t) \approx 0,\ \pi_2'(e) < 0,\ \pi_2''(e) > 0,$$

where $\pi \in [0, 1]$ *and* $e \in \mathbb{R}^+$*.*

The intuition behind this assumption is that within a policy-year, it may be rather expensive or even impossible to reduce the disease probability. However, long-term prevention may be achieved with only little effort in the preceding periods. For example: The decision to give up smoking will reduce your cancer risk after many years, only. So, for the sake of simplicity, we analyze the case when only the size of e is of choice. Furthermore, we set $\pi_1 = 1$ and denote $\pi \equiv \pi_2$.

The effort to prevent diseases is supposed to reduce the insured's welfare according to the following assumption:

Assumption A. 8. $v(e) = \mathbb{E}(u_1 - e + \rho u_2)$*.*

Together with A.7, this special form will guarantee a unique optimum.

We want to compare e under a MSC and a deductible insurance when the expected indemnity payments are equal. Thus, the actuarial value of both policies is supposed to be equal. If we did not impose this condition, one could argue that the differences in prevention are only due to the different coverage. Denote a variable x with $\tilde{x}$ if $a_2 = 0$. Then, the insured's problem is

$$\max_e V = \pi v + (1 - \pi)\tilde{v} \text{ subject to } p \text{ constant.} \tag{13}$$

Proposition 4. *Suppose A.1-A.8; then, if i) there is a positive solution to problem (13), ii) the insured is sufficiently risk avers, iii) and the probability of ll is sufficiently low, the insured will take more effort to prevent future diseases under a deductible insurance than under a MSC with equal actuarial value, i.e.,* $e_{DI} > e_{MSC}$*.*

Proof. The idea of the proof is to show that an increase of the rate of saving will lower the effort to prevention. Then, clearly, e^* under a MSC has to be lower than under a DI for all d. The first order condition requires that

$$V_e' = \pi_e'(v - \tilde{v}) - 1 \stackrel{!}{=} 0, \tag{14}$$

[10] Single-period models of the moral hazard problem and insurance can be found, for instance, in Breyer and Zweifel (1997), Drèze (2002), Shavell (1979), and Winter (1992).

which is met due to condition i) of proposition 4. Applying the implicit function theorem leads to

$$\frac{\partial e^*}{\partial \delta} = -\frac{V''_{e\delta}}{V''_{ee}}. \tag{15}$$

The second order condition requires that

$$V''_{ee} = \pi''_{ee}(v - \tilde{v}) < 0, \tag{16}$$

which is true due to A.7 and the fact that $v < \tilde{v}$ per definition. Thus, the sign of (15) is determined by the sign of

$$V''_{e\delta} = \pi'_e(v'_\delta - \tilde{v}'_\delta). \tag{17}$$

Due to A.7 $\partial e^*/\partial \delta < 0$ if and only if $v'_\delta - \tilde{v}'_\delta > 0$. Note that the marginal utilities in (17) differ only in the second period. Let $F(ll) = \theta_{ll}$ and so forth. Hence,

$$\begin{aligned} v'_\delta - \tilde{v}'_\delta = \rho\{&\mathbb{E}_{ll}[u'_2(r(d-a_1) + r\delta d')] - d'u'_d(\theta_{lh} + \theta_{hh}) \\ &- \mathbb{E}_{ll}[\tilde{u}'_2(r(d-a_1) + r\delta d')] - \mathbb{E}_{lh}[\tilde{u}'_2(r(d-a_1) + r\delta d')]\}. \end{aligned} \tag{18}$$

Because $\rho > 0$, we can safely ignore it. Dividing by u'_d and setting $\beta(a_1) = u'_2/u'_d$ leads to

$$\begin{aligned} &\mathbb{E}_{ll}[(\beta - \tilde{\beta})(r(d-a_1) + r\delta d')] - d'(\theta_{lh} + \theta_{hh}) \\ &- \mathbb{E}_{lh}[\tilde{\beta}(r(d-a_1) + r\delta d')] \lesseqgtr 0. \end{aligned} \tag{19}$$

Note that $d' < 0$ is the reaction of the MOP when δ increases due to the constant premium. In appendix A.4, we will calculate d'. Denote the first summand of (19) with A. Then, dividing by $-d'$ and applying the mean value theorem of integration lead to

$$\frac{A}{(-d')} + \theta_{hh} + (1 + \tilde{\beta}(\alpha) r\delta)\theta_{lh} \lesseqgtr \tilde{\beta}(\alpha')\{(1 + r\delta)\theta_{lh} + r/\pi(\theta_{hl} + \theta_{hh}) + \theta_{hh}\}, \tag{20}$$

where $\alpha, \alpha' \in [0, d]$ must not be equal. Now, to proof that $\partial e^*/\partial \delta < 0$, we have to show that the lhs is greater than the rhs of (20). The sign of A is unknown. Hence, a sufficiently low probability of ll requires that A has no impact on the sign of $\partial e^*/\partial \delta < 0$.[11] Furthermore, due to risk aversion: $\beta < 1$. Higher risk aversion can be interpreted as a "more" concave utility function. Thus, β decreases with higher risk aversion. However, while the lhs of (20) will not be lower than $A/(-d') + \theta_{lh} + \theta_{hh}$, the rhs tends to zero. Hence, an insured is sufficiently risk avers if the lhs becomes greater than the rhs of (20). □

Before interpreting this finding, let us state the following, obvious result:

[11] One can surely find conditions for $A > 0$ but these are similarly restrictive.

Corollary 4.1. *Suppose A.1-A.8 and i) from proposition 4; then, if the insured is nearly risk neutral, she will take more effort to prevent future diseases under a MSC than under a DI with equal actuarial value, i.e.,* $e_{DI} < e_{MSC}$.

Proof. This can be immediately derived from equation (20). With a nearly risk neutral insured, β tends to one and A to zero. But, then the lhs is clearly lower than the rhs of (20). □

Under a pure MSC, the wealth in period two will be $w_2 = y - p + r(d - a_1)$ when the prevention is successful and $y - p - d$, otherwise. Hence, the difference – the expected return on prevention – is $r(d - a_1) + d$. Let D be the MOP when $\delta = 0$; then under a DI, this difference would be D. Note that $d < D < (1 + r\delta)d$ depending on a_1. Hence, corollary 4.1 states that this return is higher when setting up a MSC than under a DI. This is due to the following leveling effect: an increase of δ lowers the expected indemnity payments in period two. However, this lowers d in both periods. Hence, the d in period two does not decrease enough to equalize the expected return on prevention.

But, if the insured is sufficiently risk avers, this effect will be counteracted by decreasing marginal utility. This is the core of proposition 4: under a MSC, the welfare is always higher in period two than under a DI. Thus, a risk avers insured has more incentive to prevention under a DI. Furthermore, the requirement of a positive solution of problem (13) is obvious: if $e^*_{DI} = 0$, no further reduction of the effort is possible. In the case of ll, the incentive to prevention is low. This leads to an ambiguous impact on effort.

To sum up, if the above mentioned leveling effect is not too prominent, and if there is a real incentive to prevention, risk aversion will lead to reduced long-term prevention activities under a MSC.

In section 3.1, we have shown that a MSC is under certain circumstances superior to a DI which means that the choice of a MSC yields higher expected utility. It may be that this difference can lead to the purchase of lower coverage, which, then, can increase the incentive to prevention.

5 Conclusions

The purpose of this paper was to compare a common deductible insurance with a medical savings contract. We have shown that a medical savings contract can increase the welfare of an insurance buyer due to reduction of income risk. So, from the insurance buyer's point of view, a further spreading of MSCs would be preferable. Furthermore, under the assumption of equal costs, a pure MSC reduces the income risk also more than a private savings plan. Hence, the insurer should add a MSA-option instead of leaving the insured on her own.

The idea that a MSC will lead always to more disease prevention activities is not as unambiguous as one might think. The key prerequisite of a comparison of insurance policies is the equal size of the actuarial value of the policies. Supposing that, however, leads to ambiguous results, i.e., the possibility of a decrease in effort to prevention. Hence, if a reform of a country's health care financing system aims to reduce the moral hazard problem with the help of an introduction of MSCs, one has to impose an increase of the insured's contribution to her medical expenses. Otherwise, one cannot be sure that the effort to prevention will actually increase.

A The Proofs

A.1 Proof of Lemma 2

With $\delta = 0$ and $v'_d \geq 0$

$$-p'_d v \geq \mathbb{E}_{lh}[\rho u'_2] + \mathbb{E}_{hl}[u'_1] + \mathbb{E}_{hh}[u'_1 + \rho u'_2], \tag{21}$$

where

$$-p'_d = \frac{1+\lambda}{1+r} \left(\mathbb{E}_{lh}[1] + \mathbb{E}_{hl}[r] + \mathbb{E}_{hh}[1+r] \right). \tag{22}$$

Setting (22) into (21), multiplying both sides by r and dividing by the term in the brackets of (22), we get

$$r\frac{1+\lambda}{1+r} v' \geq \frac{\mathbb{E}_{lh} r\rho u'_2 + \mathbb{E}_{hl} r u'_1 + \mathbb{E}_{hh}[r u'_1 + r\rho u'_2]}{\mathbb{E}_{lh} 1 + \mathbb{E}_{hl} r + \mathbb{E}_{hh}(1+r)}. \tag{23}$$

Because of $u'_{1,2} = u'(y - p - d)$ for every u' on the rhs of (23), we can place it outside the brackets. Simplifying yields the result of the lemma.

A.2 Proof of Corollary 1.1.

At the point $\delta = 0$ the arguments of the utility function are identical in both periods: income minus premium and medical costs. If the medical costs are independent and identical distributed, the difference between the expected marginal utilities without the multiplier $(d - a_1)$ becomes zero. However, the multiplier $(d - a_1)$ is high exactly in those situations where the medical costs are low in the first period. Therefore, $(d - a_1)$ weights those cases more where the difference of the marginal utilities is negative.

Proof. With $r\rho = 1$, $\mathbb{E}_{gg}((d - a_1)(u'_1 - u'_2))$ correspondents at the point $\delta = 0$ to:

$$\int_0^d \int_0^d (d - a_1)(u'_1 - u'_2) f(a_1) f(a_2) da_2 da_1 \tag{24}$$

with $u'_1 = u(Y - P - a_1)$ and $u'_2 = u(Y - P - a_2)$. Cutting the intervals $[0, d]$ of a_1 and a_2 into $[0, d/2]$ and $(d/2, d]$ leads to

$$\begin{aligned} &\int_0^{d/2}\int_0^{d/2}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1 \\ &+\int_{d/2}^{d}\int_{d/2}^{d}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1 \\ &+\int_0^{d/2}\int_{d/2}^{d}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1 \\ &+\int_{d/2}^{d}\int_0^{d/2}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1. \end{aligned} \tag{25}$$

The signs of the first two summands are unknown until now. However, one can see that the third summand is negative and the fourth positive because of the bisection. Furthermore, the assumption of independent and identical distributed medical costs ensures that the following applies:

$$\begin{aligned} &\left|\int_0^{d/2}\int_{d/2}^{d}(u'_1-u'_2)f(a_1)f(a_2)da_2da_1\right| = \\ &\qquad\left|\int_{d/2}^{d}\int_0^{d/2}(u'_1-u'_2)f(a_1)f(a_2)da_2da_1.\right| \end{aligned} \tag{26}$$

Multiplying both sides of (26) with $(d - \frac{d}{2})$ and recognizing that

$$\begin{aligned} &\left|\int_0^{d/2}\int_{d/2}^{d}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1\right| \\ &> \left|\int_0^{d/2}\int_{d/2}^{d}(d-\frac{d}{2})(u'_1-u'_2)f(a_1)f(a_2)da_2da_1\right| \end{aligned} \tag{27}$$

and

$$\begin{aligned} &\left|\int_{d/2}^{d}\int_0^{d/2}(d-a_1)(u'_1-u'_2)f(a_1)f(a_2)da_2da_1\right| \\ &< \left|\int_{d/2}^{d}\int_0^{d/2}(d-\frac{d}{2})(u'_1-u'_2)f(a_1)f(a_2)da_2da_1,\right| \end{aligned} \tag{28}$$

applies, leads to the conclusion that the absolute value of the third summand in (25) is higher than the value of the forth summand. Therefore, the sum of both is negative. The first two summands could be divided accordingly, and the argumentation is the same as before. Again, one gets a sum of two terms which is negative and two terms with unknown sign. This halving could be done again and again, so that the overall sum is negative. Therefore, the condition for a positive rate of saving from lemma 3 is met. □

A.3 Proof of Proposition 2

Define $\mathcal{D} = \{(d,\delta) : \; V_d' \geq 0 \;\&\; \delta > 0\}$. For all $(d,\delta) \in \mathcal{D}$ equation (23) becomes

$$r\frac{1+\lambda}{1+r}v' \geq \frac{r\delta\{\mathbb{E}_{ll}[u_1' - u_2'] + \mathbb{E}_{lh}[u_1' - u_d']\}}{\mathbb{E}_{lh}(1+r\delta) + \mathbb{E}_{hl}r + \mathbb{E}_{hh}(1+r)} + u_d'. \tag{29}$$

Setting (29) into (10) leads to

$$\begin{aligned} v_\delta' \geq &- \mathbb{E}_{ll}[(d-a_1)(u_1' - u_2')] \\ &+ \mathbb{E}_{lh}\left\{(d-a_1)\left[\frac{r\delta\{\mathbb{E}_{ll}[u_1' - u_2'] + \mathbb{E}_{lh}[u_1' - u_d']\}}{\mathbb{E}_{lh}(1+r\delta) + \mathbb{E}_{hl}r + \mathbb{E}_{hh}(1+r)} + u_d' - u_1'\right]\right\}. \end{aligned} \tag{30}$$

Therefore, to get $v_\delta' > 0$ for all $(d,\delta) \in \mathcal{D}$ we need that

i) the second summand of (30) is positive,

ii) and the rhs of (30) is positive, too.

If i) holds, then, clearly, ii) can be ever established by increasing $\phi \equiv F(lh)/F(ll)$. Let ϕ_1 be the lowest ϕ where rhs is positive for all $(d,\delta) \in \mathcal{D}$. So, we only have to show that i) is true.

The second summand of (30) is positive whenever

$$r\delta\{\mathbb{E}_{ll}[u_2' - u_1'] + \mathbb{E}_{lh}[u_d' - u_1']\} < (u_d' - u_1')\{\mathbb{E}_{lh}(1+r\delta) + \mathbb{E}_{hl}r + \mathbb{E}_{hh}(1+r)\} \tag{31}$$

On the lhs of (31), the first summand in the brackets should be negative because of the transfer of money from period one to two. Then, due to $\mathbb{E}_{lh}[u_d' - u_1'] \leq (u_d' - u_1')$, condition (31) is satisfied whenever

$$r\delta < \mathbb{E}_{lh}(1+r\delta) + \mathbb{E}_{hl}r + \mathbb{E}_{hh}(1+r). \tag{32}$$

However, if $\mathbb{E}_{ll}[u_2' - u_1'] > 0$, we have to increase the required ϕ until $\mathbb{E}_{ll}[u_2' - u_1'] + \mathbb{E}_{lh}[u_d' - u_1'] \leq (u_d' - u_1')$; and condition (32) can be applied. Let ϕ_2 be the lowest ϕ where this is true for all $(d,\delta) \in \mathcal{D}$.

Substituting 1 by $\mathbb{E}_{ll}1 + \mathbb{E}_{lh}1 + \mathbb{E}_{hl}1 + \mathbb{E}_{hh}1$ in (32) leads to

$$\mathbb{E}_{ll}[r\delta] - \mathbb{E}_{lh}[1] < \mathbb{E}_{hl}[r(1-\delta)] + \mathbb{E}_{hh}[r(1-\delta)+1]. \tag{33}$$

The rhs of (33) is non-negative. Let ϕ_3 be the lowest ϕ where the lhs of (33) becomes negative for all $(d,\delta) \in \mathcal{D}$.

To sum up, $v_\delta' > 0$ for all $(d,\delta) \in \mathcal{D}$ if the joint distribution function exhibits a $\phi \geq \max\{\phi_1, \phi_2, \phi_3\}$.

A.4 The Impact of δ on d

Because the premium is fixed an increase of δ has to be absorbed by a decreasing d. The premium function is

$$p(d, \delta) = \frac{1+\lambda}{1+r}\mathbb{E}(ri_1 + \pi i_2). \tag{34}$$

Applying the implicit function theorem gives

$$d' \equiv \frac{\partial d}{\partial \delta} = -\frac{p'_\delta}{p'_d} \tag{35}$$

which is equivalent to

$$\frac{\partial d}{\partial \delta} = -\frac{\pi \mathbb{E}_{lh}[r(d - a_1)]}{\pi \mathbb{E}_{lh}[1 + r\delta] + \mathbb{E}_{hl}[r] + \mathbb{E}_{hh}[\pi + r]}. \tag{36}$$

References

American Academy of Actuaries (1995a). Medical savings accounts: An analysis of the Family Medical Savings and Investment Act of 1995. Public policy monograph, American Academy of Actuaries, Washington, DC.

American Academy of Actuaries (1995b, May). Medical savings accounts: Cost implications and design issues. Public Policy Monograph 1, American Academy of Actuaries, Washington, DC.

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review 53*(5), 941–973.

Barr, M. (2001). Medical savings accounts in singapore: A critical inquiry. *Journal of Health Politics, Policy and Law 26*(4), 709–726.

Breyer, F. and P. Zweifel (1997). *Gesundheitsökonomie* (2. ed.). Springer.

Bunce, V. (2001, August). Medical savings accounts: Progress and problems under HIPAA. Policy Analysis 411, Cato Institut, Washington, DC.

Cohen, A. (2001). Per-loss vs. aggregate deductibles in insurance contracts.

Drèze, J. (2002). Loss reduction and implicit deductibles. Core Discussion Paper 5, Center of operations research and econometrics.

Eeckhoudt, L., C. Gollier, and H. Schlesinger (1991). Increases in risk and deductible insurance. *Journal of Economic Theory 55*, 435–440.

Feldstein, M. and D. Altman (1998). Unemployment insurance savings accounts. NBER Working Papers 6860, National Bureau of Economic Research.

Friedman, M. (2001, Winter). How to cure health care. *The Public Interest* (142), 3–30.

General Accounting Office (1998, December). Medical savings accounts: Results from surveys of insurers. Report to Congressional Committees GAO/HEHS-99-34, United States General Accounting Office, Washington, DC.

Gould, J. (1969). The expected utility hypothesis and the selection of optimal deductibles for a given insurance policy. *The Journal of Business 42*(2), 143–151.

Heffley, D. and T. Miceli (1998). The economics of incentive-based health care plans. *The Journal of Risk and Insurance 65*(3), 445–465.

Internal Revenue Service (2001). Medical savings accounts. Publication 969, Internal Revenue Service.

Klein, R. (2002). Optimal Deposits with a Flexible Medical Savings Account. unpublished discussion paper.

Massaro, T. and Y. Wong (1995). Positive experience with medical savings accounts in Singapore. *Health Affairs 14*(2), 267–272.

Massaro, T. and Y. Wong (1996, April). Medical savings accounts: The Singapore experience. NCPA Policy Report 203, National Center for Policy Analysis, Dallas.

Matisonn, S. (2000, June). Medical savings accounts in South Africa. NCPA Policy Report 234, National Center for Policy Analysis, Dallas.

Moon, M., L. M. Nichols, and S. Wall (1996, March). Medical savings accounts: A policy analysis. Technical report, Urban Institut, Washington, DC.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy 76*, 553–568.

Orszag, M. and D. Snower (2002). From unemployment benefits to unemployment accounts. IZA Discussion Paper 532, Institute for the Study of Labor.

Ozanne, L. (1996). How will medical saving accounts affect medical spending? *Inquiry 33*(3), 225–235.

Pauly, M. and J. Goodman (1995). Tax credits for health insurance and medical savings accounts. *Health Affairs 14*(1), 126–139.

Pauly, M. and B. Herring (2000). An efficient employer strategy for dealing with adverse selection in multi-plan offerings: An MSA example. *Journal of Health Econonomics 19*, 513–528.

Scandlen, G. (1998, July). Medical savings accounts: Obstacles to their growth and ways to improve them. NCPA Policy Report 216, National Center for Policy Analysis, Dallas.

Schlesinger, H. (1981). The optimal level of deductibility in insurance contracts. *The Journal of Risk and Insurance 48*, 465–481.

Shavell, S. (1979). On moral hazard and insurance. *Quarterly Journal of Economics 93*, 541–562.

Stano, M. (1981). Individual health accounts: An alternative health care financing approach. *Health Care Financing Review 3*(1), 117–125.

Winter, R. (1992). Moral hazard and insurance contracts. In G. Dionne (Ed.), *Contributions to Insurance Economics*, pp. 61–96. Kluwer.

Yip, W. and W. Hsiao (1997). Medical savings accounts: Lessons from china. *Health Affairs 16*(6), 244–251.

Zabinski, D., T. M. Selden, J. F. Moeller, and J. S. Banthin (1999). Medical saving accounts: Microsimulation results from a model with adverse selection. *Journal of Health Econonomics 18*, 195–218.