Department of Economics Working Paper Series

The Economics of Incentive-Based Health Care Plans

Dennis Heffley
University of Connecticut

Thomas J. Miceli
University of Connecticut

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

Abstract

Medical savings accounts (MSAs) belong to a larger class of incentive-based health care plans. Using a model that allows the consumer to invest in healthy activities, we examine the efficiency properties of incentive plans and compare them to traditional full- coverage and deductible plans, under both experience rating and community rating. The model also is extended to include utilization of preventive health care. Properly constructed incentive plans have the capacity to induce socially efficient levels of healthy activities and preventive care, raising the expected wealth of consumers without reducing insurers' profits.

The Economics of Incentive-Based Health Care Plans

Dennis Heffley and Thomas Miceli

Department of Economics, University of Connecticut

July 13, 1997

ABSTRACT

Medical savings accounts (MSAs) belong to a larger class of incentive-based health care plans. Using a model that allows the consumer to invest in healthy activities, we examine the efficiency properties of incentive plans and compare them to traditional full-coverage and deductible plans, under both experience rating and community rating. The model also is extended to include utilization of preventive health care. Properly constructed incentive plans have the capacity to induce socially efficient levels of healthy activities and preventive care, raising the expected wealth of consumers without reducing insurers' profits.

1. Background.

Comprehensive plans to overhaul the health care system have given way to "incremental" reforms, such as those designed to increase the portability of employment-based health insurance and protection for persons with pre-existing conditions. A major sticking-point in the debate of incremental reforms has been the effort to expand the availability of medical savings accounts (MSAs).[1] MSAs exist in various forms, mostly within self-insured firms, but certain features are common. MSA plans typically replace more comprehensive coverage with catastrophic or high-deductible coverage, coupled with stronger personal incentives to prudently utilize routine care and to adopt a healthier lifestyle. Consequently, MSAs can be viewed as one example of a larger class of incentive-based health insurance plans.

Incentive plans are not new. An interesting U.S. example is the Mendocino (or "Stay-Well") Plan that was adopted by the Mendocino County Office of Education in the late 1970s. Unhappy with rising health insurance premiums and facing the revenue pressure of Proposition 13, a California property tax limitation, public school officials in Mendocino County replaced their employees' first-dollar coverage Blue Shield plan with a $500-deductible policy from the same insurer. Premium savings of almost $480 per worker were used by the employer to establish a "side-fund." The fund was used to self-insure the first $500 of health care bills for each employee; amounts beyond this were still covered by Blue Shield. Apart from the change in payer for the

---

[1]. Pauly and Goodman (1995) have offered a specific MSA proposal that has drawn some lively response. See *Health Affairs*, 14 (Summer 1995), 260-279. Newhouse (1996) also offers a concise description of a voluntary MSA option for Medicare beneficiaries. Congressional passage of the 1996 Kassebaum-Kennedy health care bill (Health Insurance Portability and Accountability Act) authorized a four-year test of MSAs, limiting the number of

first $500 of claims, employees were unaffected, but an interesting twist was added -- one that reportedly influenced health care use and other forms of employee behavior.

Employees were informed that any portion of the $500 in the side-fund not used to cover a worker's initial claims would be received by that person upon quitting or retiring. A worker with only $200 in claims during the year, for example, would receive a $300 credit. Yearly credits would accumulate until voluntary severance. This incentive provision had several potential effects. First, the rebate for low utilization allowed the individual to convert some "unused" insurance into a future cash payment. Second, since the size of the rebate diminished with utilization, the employee now faced a shadow price (reduction in future bonus) for initial units of care, even though this care was fully covered by the side-fund. Third, beyond rewarding workers for monitoring their own utilization and providers' charges, it was thought that the incentive structure might encourage healthier lifestyles that would reduce the need for care and allow workers to enjoy better health and increased productivity, as well as future cash benefits.

Reactions to the Mendocino Plan and other incentive-based health plans have been predictably mixed. Employers and employees who participate in such programs often express satisfaction with the results. Insurers, especially those not offering incentive plans, usually resist such approaches, perhaps because they often entail lower-premium policies. Finally, some health care providers and outside observers warn that incentive plans may deter preventive care and raise overall health care costs. Given the structure and potential effects of such plans, it is easy to understand the conflicting views. It is more difficult, given the scarcity of empirical or theoretical research on such plans,[2] to assess the actual merits or hazards of incentive-based health insurance. The purpose of this paper is to offer a simple theoretical framework to address specific economic questions raised by this class of insurance plans: How do incentive plans influence consumer behavior, particularly various forms of illness prevention, relative to behavior in traditional plans? Can incentive plans benefit consumers and thereby lure them away from traditional plans? Are

---

tax-advantaged MSAs to 750,000 small-business or self-employed policyholders. Swartz (1996) argues strongly for researchers' input into the design and monitoring of this demonstration project.

[2] 2. Zweifel (1992) analyzes the effects of German incentive plans, while Bogetic and Heffley (1993) provide a detailed discussion of the Mendocino Plan and its potential application in other settings. More recently, Hsiao (1995) and Massaro and Wong (1995) offer conflicting views of Singapore's experience with MSAs. Ozanne (1996) gives some preliminary estimates of the cost-saving potential of MSAs. More theoretical studies of the efficiency of health care plans, such as Kamien and Schwartz (1973), Ellis and McGuire (1990), and Selden (1990), have considered conventional coinsurance, capitation, or prospective reimbursement plans, but have not examined the properties of MSAs or other incentive plans.

incentive plans necessarily less profitable for insurers? Do such plans result in "cream-skimming," leaving less healthy consumers facing higher costs in traditional plans?

Incentive plans may cause consumers to alter lifestyles, perhaps by engaging in healthy activities, and to adjust their use of primary or preventive care. To simplify the exposition, we first consider a model where the probability of illness is influenced only by the consumer's inherent health status and a single choice variable: the consumer's outlay on uninsured healthy activities (Section 2). Consumer behavior in the incentive plan is compared to behavior in traditional full insurance and standard deductible plans, where all three plans use experience rating based on health status. We also discuss implications for the insurer. To address concerns about potential "cream-skimming" in incentive plans, the model is modified to require community rating for two groups of consumers with distinctly different health status levels (Section 3). The question of how incentive plans affect various forms of prevention is examined more completely by expanding the original experience rating model to allow the consumer to invest in healthy activities *and* preventive health care -- both expenditures lower the probability of illness, but the first is uninsured and the latter is partially insured (Section 4). In a concluding section, we recap the key findings, discuss other extensions of the model, and raise further questions about the capacity of incentive-based insurance plans to improve efficiency in the health sector (Section 5).

2.   The Simple Model.

To compare the essential features of an incentive plan with more conventional insurance plans, we begin with a simple univariate model. Initial notation is listed below.

    y: consumer's income;
    h: consumer's outlay on uninsured healthy activities;
    s: consumer's exogenous health status (higher s indicates inherently better health);
$p(h,s)$:   probability of becoming ill; $p_h < 0$, $p_{hh} > 0$, $p_s < 0$, $p_{hs} < 0$;
    T: full cost of treatment in the event of illness (fixed);
    d: uninsured dollar cost to consumer of becoming ill (lost wages, etc.);
    E: target expenditure level under the incentive plan;
    π: insurance premium.

The consumer here is risk-neutral and maximizes expected wealth by choosing h, the outlay on healthy activities.[3] Costs represented by h may be direct (health club memberships,

---

[3]3. Risk-aversion is crucial in understanding why individuals voluntarily pay more than actuarially fair premiums to purchase insurance, or why some purchase more insurance than others. It is less critical in

exercise equipment, etc.) or indirect, such as the opportunity cost of time used to exercise or prepare healthier meals.[4] We assume that h is uninsurable and therefore different from preventive health care, which often is partially insured; the latter is introduced in Section 3. The inherent (exogenous) health status of the consumer is denoted by s, where higher values of s mean better health. We assume that the consumer knows s, which also can be observed by the insurer, based on age, health history, physical examinations, and the like. This observability of health status allows the insurance contract to be conditioned on s, as in the case of experience rating.[5] We assume that a contract cannot be conditioned on h, the consumer's investment in healthy activities.

The assumption that $p_{hs} < 0$ implies that the marginal productivity of healthy activities is higher for an inherently healthier person. Someone with chronic back problems, for example, may benefit little from joining a health club. This may be the more plausible assumption, but the reverse clearly could hold for some individuals or some types of chronic conditions. Where appropriate, we note how the assumption affects certain results.

The socially efficient level of healthy activities for a person of health status s, denoted h*(s) or simply h*, minimizes the cost of such activities plus the expected social cost of being ill, or

$$h + p(h,s)(d+T), \tag{1}$$

which yields the first-order condition

---

understanding the response of insured individuals to the structure of the plan. Moreover, the assumption of risk-neutrality in the present context is not too restrictive because the consumer will have full health insurance against the cost of treatment in the event of illness under both the traditional and incentive plans. Landes and Posner (1987, pp. 55-58) discuss this assumption in the context of liability rules in tort law.

Also note that the present U.S. tax exemption of health insurance premiums paid by employers and employees offers large subsidies to the purchase of health insurance. Under such conditions, individuals who are risk-neutral (or even risk-seeking) may rationally insure. Various proposals to make health insurance compulsory (employer mandates or individual mandates) also suggest that risk-aversion is not a necessary condition for insurance in a political economy.

[4]4. We thank a referee for pointing out that our analysis of consumer behavior aimed at reducing the likelihood of illness is similar to Ehrlich and Becker's (1972) analysis of "self-protection" as an alternative to market insurance. As in our model, their analysis shows that incentives to engage in self-protection are not solely derived from risk-aversion and will exist in risk-neutral individuals as well. Also consistent with our model, they show that full insurance creates a moral hazard problem that reduces the incentive of consumers to engage in self-protection. However, this disincentive can be offset if the insurance premium is decreasing in the amount of self-protection (Ehrlich and Becker, 1972: pp. 640-642). In our model, we assume that this effect is absent since most forms of healthy activity are difficult for insurers to monitor (i.e., we assume $\partial \pi / \partial h = 0$).

[5]5. We do not consider the implications of risk factors known only to the consumer, since our concern here is with moral hazard rather than adverse selection problems. On the latter, see Rothschild and Stiglitz (1976) or Van de Ven and Van Vliet (1995).

$$1 + p_h(h^*,s)\,(d+T) \;=\; 0. \tag{2}$$

The second-order condition for a minimum holds, given $p_{hh} > 0$. Implicit differentiation of (2) shows that a healthier individual optimally invests more in healthy activities ($\partial h^*/\partial s = -p_{hs}/p_{hh} > 0$) when $p_{hs} < 0$.

2.1 *Full Insurance Plan*

When the consumer has no insurance and bears the full cost of h, d, and T, individual maximization of expected wealth, $EW = y-h-p(h,s)(d+T)$, satisfies (2) and therefore also minimizes social costs. On the other hand, if the consumer pays a fixed premium, π, to fully insure for the cost of treating an illness, T, expected wealth is given by [6]

$$EW = y - h - p(h,s)d - \pi. \tag{3}$$

Here the fully insured consumer's choice of h, denoted $\underline{h}(s)$ or simply $\underline{h}$, solves

$$1 + p_h(\underline{h},s)d \;=\; 0. \tag{4}$$

Because of the effect of s on the marginal benefit of healthy activities, $p_h(\underline{h},s)d$, the consumer's choice of $\underline{h}$ under full insurance is sensitive to health status, but does not depend on whether or not the premium is conditioned on s, since the consumer takes both π and s as given. Comparing (2) and (4) shows that $\underline{h} < h^*$ for all s.[7] Regardless of health status level, the consumer invests too little in healthy activities from a social perspective. This reflects the standard moral hazard problem associated with full insurance, which causes the consumer to ignore the expected costs of treatment, $p(h,s)T$.

In this section, we focus on experience rated insurance contracts, where the premium is conditioned on s. The expected return for the insurance company under this plan is given by [8]

$$\pi(s) - p(\underline{h},s)T, \tag{5}$$

where $\pi(s)$ is the experience rated premium and $p(\underline{h},s)T$ is the expected cost of health care when the fully insured consumer chooses $\underline{h}$. If competition drives insurers' profits to zero, $\pi(s) = p(\underline{h},s)T$ in equilibrium for all s. Differentiating with respect to s shows that

$$\pi_s = p_s T + p_h(\partial \underline{h}/\partial s)T < 0. \tag{6}$$

Healthier consumers face lower premiums due to lower inherent risk of illness (first term) and because they invest more in healthy activities (second term). Had we assumed that sicker persons

---

[6]6.   We assume that the cost d is not insured.

[7]7.   Equating the two conditions and rearranging terms gives: $p_h(\underline{h},s)/p_h(h^*,s) = (d+T)/d > 1$ or, since $p_h < 0$, it follows that $p_h(\underline{h},s) < p_h(h^*,s)$. This implies that $\underline{h} < h^*$, since $p_{hh} > 0$.

[8]8.   For simplicity, we assume that the insurer's administrative costs are zero here and in the other plans considered.

have a higher marginal productivity of h (i.e., $p_{hs} > 0$), the second effect would be positive and the overall effect of s on the premium would be ambiguous.

Substituting the equilibrium premium, $\pi(s) = p(\underline{h},s)T$, into (3), the consumer's expected wealth, conditional on s, becomes

$$\underline{EW}(s) = y - \underline{h} - p(\underline{h},s)(d+T). \tag{7}$$

This is necessarily less than expected wealth if the consumer had chosen h* since, by definition, h* minimizes $h+p(h,s)(d+T)$ and therefore maximizes $EW = y-h-p(h,s)(d+T)$. Consumers ultimately bear the inefficiency associated with moral hazard if insurance markets are sufficiently competitive; but, as shown below, the insurance plan can be restructured to improve this outcome.

2.2   *Incentive Plan*

Consider an alternative plan that provides full insurance for treatment costs, but which also stipulates that if the consumer's health care spending during a given period falls below a target level, the difference between the target and the actual spending is refunded to the consumer. As above, we assume that the terms of this plan can be conditioned on s. Specifically, the consumer receives full coverage for his treatment costs, plus an expected refund of

$$R(s) = E(s) - p(h,s)T, \quad \text{if } p(h,s)T < E(s),$$
$$R(s) = 0, \quad\quad\quad\quad \text{if } p(h,s)T \geq E(s), \tag{8}$$

where E(s) is the expenditure target for consumers of health status s. More compactly, the expected refund is $R(s) = \max\{E(s)-p(h,s)T, 0\}$.

The consumer's expected wealth under the incentive plan is

$$EW(s) = y - h - p(h,s)d + R(s) - \pi(s), \tag{9}$$

where, for now, we assume that $\pi(s)$ is the same premium as in the traditional plan. What choice of h by the consumer maximizes this expression? If $R(s) = 0$, that is if $p(h,s)T \geq E(s)$, the expression is identical to (3) and the optimal choice is $\underline{h}$. Consequently, if the expected refund is zero, the consumer simply acts as if the plan were a traditional full insurance plan. In contrast, if the expected refund is positive, then substituting $R(s) = E(s)-p(h,s)T$ into (9) gives $EW(s) = y-h-p(h,s)(d+T)+E(s)-\pi(s)$. The first-order condition for maximization of this expression is the same as (2), so the incentive plan consumer who anticipates a positive refund will choose h*, the socially efficient level of healthy activities for his particular health status (recall that h* varies with s). The incentive plan causes the consumer to internalize the impact of his healthy activities on the expected costs of treatment [i.e., the effect of h on $p(h,s)T$] through the refund.

Whether the consumer chooses h* or h̲ depends on the size of E(s). In general, the larger is E(s), the more likely the refund will be positive, and the more likely the consumer will prefer h* to h̲. Figure 1 illustrates a case where the consumer selects h*, and Figure 2 illustrates a case where he chooses h̲. Both figures graph the expressions y-h-p(h,s)d-π(s) [maximized at h̲] and y-h-p(h,s)(d+T)+E(s)-π(s) [maximized at h*]. The first expression is relevant when p(h,s)T ≥ E(s) and the refund is zero; the second is relevant when p(h,s)T < E(s) and the refund is positive. The transition point occurs where the curves intersect, since p(h,s)T = E(s) at that point. This point is labelled h_o; when $h < h_o$ the leftmost curve is relevant, and when $h > h_o$, the rightmost curve is relevant. The upper envelope of the two curves is the relevant surface. The amount of the refund under the incentive plan is given by the vertical distance between the two curves to the right of h_o (see Figure 1).

[Figures 1 and 2]

The consumer's choice of h in the incentive plan is found by comparing the peaks of the two curves. In Figure 1, h* is preferred because

$$[y - h^* - p(h^*,s)(d+T) + E(s) - \pi(s)] > [y - \underline{h} - p(\underline{h},s)d - \pi(s)]. \tag{10}$$

Here the incentive plan leads to a socially efficient level of healthy activities, h*, and reduces p(h,s)T to a level where the consumer receives a rebate of E(s)-p(h*,s)T. On the other hand, in Figure 2, the above inequality is reversed and the incentive plan member chooses h̲, receiving no rebate and behaving the same as a fully insured consumer with the same health status.

As E(s) is increased, the left-hand side of (10) increases (i.e., the rightmost curve in the graphs shifts up). Thus, there exists a critical value of E(s), denoted $E^c(s)$, such that for E(s) < $E^c(s)$ the consumer chooses h̲, and for E(s) ≥ $E^c(s)$ the consumer chooses h*.[9] Figure 3 shows the consumer's choice of h for various levels of E(s). The expression for the critical target expenditure level for a consumer of health status s, denoted $E^c(s)$, is found by evaluating each curve at its peak, equating the expressions, and solving for E:

$$E^c(s) = p(h^*,s)T + [p(h^*,s)d+h^*] - [p(\underline{h},s)d+\underline{h}]. \tag{11}$$

$E^c(s)$ is strictly positive given that $[p(h^*,s)d+h^*]-[p(\underline{h},s)d+\underline{h}] > 0$, by definition of h̲ as the value of h that maximizes (3) and, hence, minimizes [p(h,s)d+h]. Thus, *some* expenditure target is required to induce socially efficient behavior. Also, by (11), the rebate at the consumer's point of indifference between the two plans, $E^c(s)-p(h^*,s)T$, is strictly positive.[10]

---

[9]9. We assume that when indifferent the consumer chooses h*.

[10]10. In Figure 1, it is easy to see that when the target is high enough that the two curves attain the same height, thereby defining $E^c(s)$, the rebate at h* will remain strictly positive.

[Figure 3]

From (11) we can determine how $E^c(s)$ varies with s:[11]

$$\partial E^c/\partial s = p_s(h^*,s)T + [p_s(h^*,s)-p_s(\underline{h},s)]d < 0, \qquad (12)$$

where the expression in brackets is negative by $p_{hs} < 0$, given that $h^* > \underline{h}$. Thus, the critical target spending level needed to draw consumers into the incentive plan and to induce efficient behavior is lower for healthier consumers.[12] This result shows that even less healthy consumers can be drawn into an incentive plan and induced to choose an efficient level of healthy activities, h^*, provided the target spending level is set sufficiently high. Relaxing the spending target by setting it higher gives even consumers with poorer health an opportunity to benefit from the refund incentive by investing more in healthy activities and reducing the probability of illness. This contradicts the view that incentive plans inevitably lead to "cream-skimming" and the isolation of less healthy consumers in traditional plans.

In addition to varying with a consumer's health status, $E^c(s)$ is sensitive to the fixed cost of treatment [$\partial E^c/\partial T = p(h^*,s) > 0$], the consumer's uninsured dollar cost of becoming ill [$\partial E^c/\partial d = p(h^*,s)-p(\underline{h},s) < 0$], and the technology of prevention, $p(h,s)$. This sensitivity argues strongly for the need to adjust the expenditure target as underlying conditions of the medical marketplace change.[13]

The preceding analysis shows that it is possible to induce consumers of varying health status to choose efficient levels of healthy activities under the incentive plan. The next question is whether the insurer also has an incentive to offer this plan. Does the incentive plan increase the insurer's return relative to the traditional plan? Recall that (5) gives the insurer's expected return under the traditional plan. Assuming the same premium $\pi(s)$, and retaining the assumption of experience rating, the expected return from insuring a person of health status s under the incentive plan is the premium less the expected cost of treatment and the expected rebate, or $\pi(s)-p(h,s)T-$max$\{E(s)-p(h,s)T, 0\}$. Given optimal consumer behavior under the incentive plan, the expected rebate must be positive [see (11) or Figure 1], so the insurer's expected return is simply $\pi(s)-E(s)$. If it offers the plan, the insurer will set E(s) as low as possible, while still inducing the consumer

---

[11]11. The Envelope Theorem was used to eliminate terms associated with $\partial h^*/\partial s$ and $\partial \underline{h}/\partial s$.

[12]12. If $p_{hs} > 0$, the impact of health status on the critical target spending would be ambiguous, but there would still exist a critical value for any observed s.

[13]13. The Mendocino Plan has retained the initial $500 expenditure target, but has adjusted other terms of the contract, including the introduction of a 20% coinsurance rate for the first $2,000 above the $500 target, giving an out-of-pocket limit of $400. Other adjustments include coverage for means of contraception and provisions for domestic partners.

to choose h*. Levels of $E(s) > E^c(s)$ yield no further increase in the consumer's healthy activities (recall Figure 3), thus the insurer will set $E(s) = E^c(s)$, yielding a return of

$$\pi(s) - E^c(s). \tag{13}$$

From (5), the expected return to the insurer offering full coverage, absent competition, was $\pi(s) - p(\underline{h},s)T$. The incentive plan will therefore provide the uncontested insurer a greater return on a person with health status s if $p(\underline{h},s)T > E^c(s)$. Using $E^c(s)$, as defined by (11), we get

$$p(\underline{h},s)T - E^c(s) = [p(\underline{h},s)(d+T) + \underline{h}] - [p(h^*,s)(d+T) + h^*] > 0, \tag{14}$$

where the inequality follows by definition of h*.[14]

In an uncontested environment, insurers earn a greater return by offering the incentive plan because it induces consumers to invest more in healthy activities. But if free entry exists, competition among insurers will cause the incentive plan premium, experience rated for each s, to fall until a zero return prevails. That is, $\pi(s) = E^c(s)$ in a competitive equilibrium. Since $\_E^c/\_s < 0$, by (12), healthy consumers would be charged less than unhealthy consumers, but they would also have lower (more restrictive) target levels of spending.

Given competition, the expected wealth of the incentive plan member with health status s is

$$EW^*(s) = y - h^* - p(h^*,s)d + [E^c(s) - p(h^*,s)T] - \pi(s) = y - h^* - p(h^*,s)(d+T). \tag{15}$$

Comparing this to the consumer's expected wealth under the traditional full coverage plan in (7) shows that $EW^*(s) > \underline{EW}(s)$. Thus, when insurance companies are competitive and spending targets are properly set, consumers at all health status levels strictly benefit from the incentive plan and insurers do no worse (zero profits). Competition forces insurers to pass the increased returns under the incentive plan to consumers in the form of lower net insurance costs. These lower net costs reflect the social gains from inducing consumers to choose efficient levels of healthy activities. If profits are not driven to zero and social gains are shared between consumers and insurers, both parties can benefit from a move from full coverage to incentive-based insurance.

2.3   *Traditional Insurance with a Deductible*

Part 2.2 compared the incentive plan to a full insurance plan that covered all health care costs. In reality, traditional insurance plans sometimes confront the moral hazard problem of too

---

[14]14. This dominance of the incentive plan requires that the insurer be able to determine $E^c(s)$. We show that $E^c(s)$ exists but, as a practical matter, determination of the critical target for any given health status is likely to require market iterations by profit-seeking insurers.

little h by imposing a deductible that requires consumers to pay the first D dollars of their health care costs, after which additional costs are partly or fully covered, depending on the presence of coinsurance. This section compares a simple deductible policy, with full coverage beyond the deductible, to the incentive plan.

Under a policy with a deductible of $D(s)$, which again depends on observed health status, the consumer's expected health care costs are

$$p(h,s)T, \qquad \text{if } p(h,s)T < D(s),$$
$$D(s), \qquad \text{if } p(h,s)T \geq D(s). \tag{16}$$

That is, the consumer's expected costs are $\min\{p(h,s)T, D(s)\}$. Expected wealth under this deductible policy is

$$EW(s) = y - h - p(h,s)d - \min\{p(h,s)T, D(s)\} - \pi(s). \tag{17}$$

Since the consumer views $D(s)$ as fixed, he will choose $\underline{h}$ if $\min\{p(h,s)T, D(s)\} = D(s)$, but he will choose h* if $\min\{p(h,s)T, D(s)\} = p(h,s)T$. The condition for the consumer to choose h* is

$$[y - h^* - p(h^*,s)(d+T) - \pi(s)] > [y - \underline{h} - p(\underline{h},s)d - D(s) - \pi(s)], \tag{18}$$

which is satisfied for a sufficiently large $D(s)$. That is, a large enough deductible will induce the consumer to choose the efficient h.

From the equality version of (18), the lowest level of $D(s)$ that accomplishes this is

$$D^c(s) = p(h^*,s)T + [p(h^*,s)d+h^*] - [p(\underline{h},s)d+\underline{h}]. \tag{19}$$

This strictly positive result is identical to (11), the expression for $E^c(s)$. Like the critical refund in the efficient incentive plan, the critical deductible forces the consumer to internalize the full expected cost of treatment. Yet the plans do differ. The consumer expects to pay the full cost of treatment under an efficient deductible plan, given that $D^c(s) > p(h^*,s)T$ by (19), though in the event of illness it is likely that $D^c(s) < T$. In contrast, all of the consumer's expected treatment costs are covered in the efficient incentive plan, and he is reimbursed for the excess of $E^c(s)$ over $p(h^*,s)T$. This is possible, in a competitive environment, because the consumer pays $\pi(s) = E^c(s)$ up front as a premium, whereas under the efficient deductible plan the consumer pays a smaller premium, but then also must pay up to $D^c(s)$ of T when illness occurs.

A useful way to think about the difference between the efficient deductible plan and the efficient incentive plan involves the notions of health risk and income risk.[15] In both plans, the consumer faces a health risk, which is realized when illness occurs. The size of this risk, $p(h^*,s)$, is the same in both plans if rebates and deductibles are properly set. The two plans differ, though, in the form and timing of income risk. Under the deductible plan, the consumer faces income risk

---

[15]15. We thank one of the referees for making this point.

in the form of uncertain out-of-pocket health care expenditures, up to the amount of the deductible, when illness occurs. Under the incentive plan, income risk takes the form of an uncertain refund of the initial premium payment at some later date. The incentive plan therefore offers the advantage of eliminating the correlation between the consumer's health risk and income risk that characterizes the deductible plan. The price of eliminating this correlation of risks is a higher premium in the incentive plan than in the deductible plan.[16]

3.   The Model with Community Rating.

We have compared the incentive plan with full insurance and with a deductible plan, assuming that all insurance contracts were experience rated for consumers of differing health status. A frequent criticism of incentive plans like MSAs, however, has been that they will attract healthier consumers away from traditional plans, thereby raising premiums for the remaining, less healthy, consumers (Nichols, 1995; Urban Institute, 1996). This argument is based on a community rating system that constrains insurers to charge enrollees the same premium regardless of health status. We examine this criticism by now comparing a traditional insurance plan and an incentive plan for the case where insurers cannot condition contracts on s.

To keep the analysis simple, consider the special case where $p_{hs} = 0$. That is, health status does not affect the marginal productivity of healthy activities; it simply affects the risk of illness. To capture this case, let $p(h,s) = p(h)-s$, where $p' < 0$, $p'' > 0$.[17] Further suppose that there are two groups, those with good health, or health status s_g, and those with poor health, s_p, where $s_g > s_p$. If α is the fraction of consumers with good health, the average health status of the population is defined as $s_a = \alpha s_g + (1-\alpha)s_p$.

3.1   *Community Rated Full Insurance*

Under full insurance with pooled risk groups, all consumers pay the same premium π, regardless of health status. A type j consumer has expected wealth of

$$EW_j = y - h - [p(h)-s_j]d - \pi, \quad j = g,p \tag{20}$$

and chooses $\underline{h}$ to solve

---

[16]16. Recall, from (13), that under competitive conditions the incentive plan premium equals $E^c(s)$. The competitive premium for the efficient deductible plan equals $p(h^*,s)[T-D^c(s)]$. But, from (11) and (19), $E^c(s) = D^c(s) = p(h^*,s)T + [p(h^*,s)d+h^*] - [p(\underline{h},s)d+\underline{h}]$. It follows that the incentive plan premium exceeds the deductible plan premium by $[p(h^*,s)]^2T + [1+p(h^*,s)]\{ [p(h^*,s)d+h^*] - [p(\underline{h},s)d+\underline{h}] \}$. This expression is strictly positive, since the final bracketed term is positive by definition of $\underline{h}$ as the value of h that maximizes (3) and, hence, minimizes $[p(h,s)d+h]$.

[17]17. This separable form for $p(h,s)$ allows us to isolate the incentive effects of insurance from its pooling effect under a community rating system.

$$1 + p'(\underline{h})d = 0, \tag{21}$$

where $\underline{h}$ is independent of s, given the assumed form of $p(h,s)$. The competitive insurance premium in this case is

$$\pi = [p(\underline{h})\text{-}s_a]T, \tag{22}$$

and the equilibrium expected wealth level of a type j consumer is

$$\underline{EW}_j = y - \underline{h} - p(\underline{h})(d+T) + s_j d + s_a T, \quad j = g,p. \tag{23}$$

Note that the final term, $s_a T$, captures the pooling of risks. Specifically, consumers in good health subsidize the treatment costs of those in poor health, since $s_g > s_a > s_p$.

3.2   *Community Rated Incentive Plan*

Now consider an incentive plan that also does not discriminate based on health status. Any consumer who chooses such a plan will pay the same premium and face the same target spending level, E, regardless of his or her health status. Under such a plan, the expected refund for a type j consumer is

$$R_j = \max\{E\text{-}[p(h)\text{-}s_j]T, 0\}, \tag{24}$$

and the consumer's expected wealth is

$$EW_j = y - h - [p(h)\text{-}s_j]d + R_j - \pi$$
$$= y - h - [p(h)\text{-}s_j](d+T) + E - \pi, \quad j = g,p \tag{25}$$

where E and π are independent of s. The choice of h that maximizes this expression, assuming a positive refund, is h*, the socially efficient level of healthy activities, which is independent of s by the separability of $p(h,s)$ and independent of the insurance contract terms E and π, which the consumer takes as given.

As before, we can derive the minimum target expenditure level that just induces a consumer to choose the incentive plan over full insurance (assuming the same premium for each):

$$E^c_j = [p(h^*)\text{-}s_j]T + [h^*\text{+}p(h^*)d] - [\underline{h}\text{+}p(\underline{h})d], \quad j = g,p \tag{26}$$

Only the first term in (26) depends on s_j, and since $s_g > s_p$, it follows that $E^c_g < E^c_p$; the minimum target level of expenditures is lower for healthier consumers, consistent with (12). Since any target level greater than E^c_j will attract a consumer of type j from the full insurance plan, it is possible to attract only healthy consumers by setting a target level between E^c_g and E^c_p, though the reverse is not possible. This is consistent with the claim that incentive plans may "skim the cream" from traditional, community rated insurance plans, increasing the cost of insurance to less healthy consumers who remain in the traditional plan. Specifically, $s_a T$ will fall to $s_p T$ in (22), raising the competitive insurance premium in the traditional plan.

"Cream-skimming" is possible, but not inevitable. Suppose that the target is set to induce both types of consumers to choose the incentive plan. The minimum target level that accomplishes this pooling is E^c_p. The insurer's expected cost under this plan is the expected cost of treatment plus the expected refund, or

$$[p(h^*)-s_a]T + \{E^c_p - [p(h^*)-s_aT]\} = E^c_p. \qquad (27)$$

Thus, the competitive premium that insurers will charge for this pooled incentive plan is $\pi = E^c_p$. Substituting this into (25) we obtain

$$EW_j^* = y - h^* - [p(h^*)-s_j](d+T), \quad j = g,p. \qquad (28)$$

Finally, we can compare the expected wealth of a consumer under the pooled traditional and pooled incentive plans by calculating the difference between (28) and (23) for both types of consumers:

$$EW_j^*-\underline{EW}_j = [\underline{h}+p(\underline{h})(d+T)] - [h^*+p(h^*)(d+T)] + (s_j-s_a)T, \quad j = g,p. \qquad (29)$$

The first two bracketed terms are on net positive and are independent of the consumer's health status.[18] This represents the efficiency gain from the incentive plan over full insurance, and is available to consumers, independent of health status. The final term is positive for consumers in good health ($s_g > s_a$), but negative for consumers in poor health ($s_p < s_a$). It represents the fact that healthier consumers cross-subsidize less healthy ones under traditional pooled insurance. This factor reinforces the efficiency gain of the incentive plan for healthy consumers and offsets the efficiency gain for less healthy ones. If the efficiency gain is strong enough, however, consumers in poor health may still enjoy a net gain under the incentive plan.

It is worth noting why the incentive plan in this section allows the healthy to avoid being pooled with consumers in poor health, even though they pay the same premium and face the same expenditure target. The key is that the *refunds* are not pooled. Although all consumers pay the same amount up-front to purchase the policy under community rating, the expected refunds of consumers in good health exceed the expected refunds of consumers in poor health [see (24)].

4. The Consumer Also Chooses Preventive Health Care.

The univariate choice models in Sections 2 and 3 capture the essential differences between the incentive plan and more conventional forms of health insurance (full coverage and deductible plans), under both experience and community rating systems. In those models the consumer's outlay on healthy activities reduces the probability of illness, conditional on his given health status,

---

[18]18. The fact that these two terms are independent of s follows from our assumption, in Section 3, that p(h,s) is separable in h and s.

but this personal outlay is not insured. Only the direct cost of treatment, T, is covered. In reality, consumers face another type of expenditure that may be partially insured and which also affects the probability of illness, namely the expenditure on preventive care, including routine medical check-ups. It is important to consider this extension, since critics of incentive plans often claim that the reward structure of such plans discourages preventive care and increases the likelihood of illness. Let the consumer's expenditures on both healthy activities, h, and preventive care, x, reduce the probability of illness, conditional on health status, or

$$p(h,x,s), \quad p_h, p_x, p_s < 0, \quad p_{hh}, p_{xx} > 0, \quad p_{hh}p_{xx} > p_{hx}^2. \tag{30}$$

Also, since we limit attention in this section to experience rated plans, the terms of insurance (premium, expenditure target, deductible) should be interpreted as functions of s.

First consider the socially efficient levels of healthy activities and preventive care. The socially efficient combination, $(h^*(s),x^*(s))$ or simply (h^*,x^*) minimizes overall expected costs, including the consumer's transaction costs, given by

$$h + x(1+t) + p(h,x,s)(d+T), \tag{31}$$

where t represents the transaction cost of consuming preventive health care, expressed as a fraction of preventive expenditures. We introduce t to capture the full social cost of preventive care, which includes not only the full charges, x, which may be partially insured, but also the time, travel, or other indirect costs of consuming preventive care that are not insurable. This distinction did not arise with h, since both the direct and indirect costs of engaging in healthy activities were assumed to be uninsurable. The resulting first-order conditions are

$$1 + p_h(h^*,x^*,s)(d+T) = 0, \tag{32}$$

coinsurance rate c for the direct dollar costs of prevention ($0 \le c \le 1$). The consumer's expected wealth under this plan is

$$EW = y - h - p(h,x,s)d - (c+t)x - \pi, \tag{34}$$

where the consumer again takes the insurance premium, π, as given. The consumer's choices of h and x, denoted $(\underline{h}(s),\underline{x}(s))$ or simply $(\underline{h},\underline{x})$, satisfy the conditions

$$1 + p_h(\underline{h},\underline{x},s)d = 0, \tag{35}$$
$$(c+t) + p_x(\underline{h},\underline{x},s)d = 0. \tag{36}$$

Sign restrictions in (30) ensure that $(\underline{h},\underline{x})$ is a maximum. Condition (35) is the same as (4), and again suggests that, from a social perspective, the consumer underinvests in healthy activities in the traditional plan. The latter condition (36) contains two sources of inefficiency. First, given full insurance, the consumer ignores the cost of treatment, T. This tends to lower x below x*. On the other hand, if the consumer pays less than the full cost of prevention (c < 1), he will tend to

-14-

consume more than x*. Note that the tendency to underconsume x is due to full insurance for treatment charges, whereas the tendency to overconsume x is due to partial insurance for the costs of prevention. In combination, these two effects can result in either too much or too little prevention. Comparing (35) and (36) to the conditions that define socially efficient behavior, (32) and (33), it is clear that consumer's choice of both h and x in the traditional plan may be distorted.

If premiums are experience rated, the expected return to the insurance company under this traditional plan is given by

$$\pi(s) - (1-c)\underline{x} - p(\underline{h},\underline{x},s)T. \tag{37}$$

In a competitive market, (37) equals zero, so the consumer's expected wealth becomes

$$\underline{EW}(s) = y - \underline{h} - p(\underline{h},\underline{x},s)(d+T) - (1+t)\underline{x}. \tag{38}$$

Now consider the incentive plan when expenditures on both healthy activities and preventive care are choice variables. The refund in (8) becomes

$$R(s) = \max\{E(s)-p(h,x,s)T-(1-c)x, 0\}. \tag{39}$$

This specification assumes that the refund is based on the *insurer's share* of expected health costs compared to the target expenditure, E(s). Thus, it depends on the rate of coinsurance, c, for preventive care. The expected wealth of the consumer under this plan is

$$EW(s) = y - h - p(h,x,s)d + \max\{E(s)-p(h,x,s)T-(1-c)x, 0\} - (c+t)x - \pi(s). \tag{40}$$

If the refund is positive, (40) becomes

$$EW(s) = y - h - p(h,x,s)(d+T) - (1+t)x - \pi(s), \tag{41}$$

which the consumer maximizes by choosing (h*,x*), since the first-order conditions for (41) are identical to (32) and (33). Proceeding as above, we can show that the consumer will make this socially efficient choice, rather than ($\underline{h},\underline{x}$), if E(s) is larger than a critical value, given by

$$E^c(s) = p(h^*,x^*,s)T + (1-c)x^* + [p(h^*,x^*,s)T + h^* + (c+t)x^*]$$
$$- [p(\underline{h},\underline{x},s)T + \underline{h} + (c+t)\underline{x}] > 0. \tag{42}$$

Zero profit for insurers implies that, in equilibrium, $\pi(s) = E^c(s)$. The equilibrium expected wealth of consumers is

$$EW^*(s) = y - h^* - p(h^*,x^*,s)(d+T) - (1+t)x^*, \tag{43}$$

which is strictly larger than $\underline{EW}(s)$ in (38).

The preceding analysis shows that the incentive plan, with $E^c(s)$ properly chosen, can induce not only the efficient level of healthy activities (as in Section 2) but also the efficient level of preventive care for consumers with varying health status levels. The question of whether the plan *reduces* the level of preventive care, as critics contend, depends on whether consumers were over- or underconsuming prevention in the traditional plan. If individuals were consuming too

little preventive care in the traditional plan, due to full coverage of treatment, a properly structured incentive plan will lead to higher, socially efficient levels of prevention. Conversely, if traditional plan consumers were overconsuming preventive care due to moral hazard effects, the appropriate incentive plan would indeed reduce prevention, but only to socially efficient levels.

5.   Summary and Possible Extensions.

Opponents of incentive-based health insurance plans emphasize two potential problems: "cream-skimming" and reductions in preventive care. This study directly addresses both issues. After recapping the point that full coverage of treatment causes underinvestment in healthy activities (Section 2.1) -- an uninsured and unobserved form of prevention -- we explore the efficiency properties of a simple incentive-based insurance plan that rewards consumers who keep the expected treatment costs paid by the insurer below some target expenditure level (Section 2.2). Results show that, under experience rating, where insurers can condition the premium and expenditure target on a person's observable health status, individuals can be induced to select socially efficient levels of healthy activities, provided that the expenditure target is set sufficiently high. Less healthy consumers require higher targets to induce incentive plan participation and socially efficient behavior, but properly constructed targets can improve the expected wealth of consumers at all health status levels, without reducing insurers' profits. This system of targets is sensitive to the costs of treatment, the dollar costs to the consumer of becoming ill, and the technology of illness-prevention, but competitive market pressures to attract consumers might prompt insurers to make many of these marginal adjustments.

We also point out the similarities between the incentive plan and a traditional deductible plan (Section 2.3). As with the introduction of an expenditure target, the deductible gives the insurer an extra optimization instrument -- an additional term of insurance. In this sense, it is not surprising that both approaches are Pareto-improving when correctly implemented. Both the rebate and the deductible, when properly structured, force the consumer to face the full expected costs of illness, but the two plans differ in the timing of the consumer's payment for this cost. Difference between these plans may be even more apparent in a less static framework where the timing of costs matters. Similarly, in a framework where consumer psychology is allowed to be more complex, carrots and sticks of equal size could have different consequences. In would be useful to explore the differences between the various types of plans in a more dynamic model that also allows for the possibility of asymmetric response to penalties and inducements.

A more dynamic framework also would lend itself to a practical aspect of incentive plans that we have not explicitly addressed: the rules governing the accumulation and distribution of savings. In particular, is the consumer's expected refund paid periodically, or is it accumulated for lump-sum distribution at a specified future date? How do consumers and insurers share the interest accrued on positive balances in the consumer's account? Deferred or partial access to accumulated savings reduces the present value of the refund. Consequently, the smaller the consumer's share of accumulated savings, the longer the wait for the reward, or the higher the rate of time preference, the larger the expenditure target would need to be to attract participation in the incentive plan and to induce socially efficient behavior.

The second major issue of "cream-skimming" is taken up in Section 3, where experience rating is replaced by community rating and the health status continuum is replaced by two groups of consumers with distinctly different health status levels. Again, full insurance is analyzed briefly and the usual cross-subsidization, under community rating, from healthy to unhealthy consumers is shown (Section 3.1). Analysis of a community rated incentive plan shows that "cream-skimming" is possible but not inevitable (Section 3.2). Setting the expenditure target just high enough to attract the healthy away from the traditional plan does result in separation of the groups. However, setting the expenditure target high enough to attract the unhealthy, and thereby inducing socially efficient behavior by this group, also will attract the healthy, who only require a lower target to join the incentive plan and to adopt efficient behavior. Provided insurance markets are competitive, both groups gain from improved incentive effects, but the removal of cross-subsidies that exist under community rated full insurance plans favors the healthy. Thus, under community rating, the healthy are unambiguously better off in the incentive plan. The unhealthy may gain or lose, depending on the size of the incentive effects relative to the size of the (eliminated) cross-subsidies. This finding does not erase concerns about "cream-skimming," but suggests that it is wrong to assume that incentive plans necessarily harm less healthy consumers.

Section 4 extends the model to provide a more complete analysis of the effects of incentive plans on prevention. In addition to investing in uninsured healthy activities, the individual consumes preventive health care, which is partially insured. Traditional insurance plans cause underinvestment in healthy activities, but may result in either too little or too much consumption of preventive health care due to two opposing effects: partial coverage of preventive care encourages its consumption, relative to the uninsured case, while more complete coverage of treatment costs reduces the incentive to consume preventive care. Analysis of the incentive plan

again shows that, for any particular health status level, the target level of expenditure can be constructed to induce selection of the incentive plan over the traditional plan and to achieve socially efficient levels of investment in both healthy activities and preventive care. If preventive care was "too low" in the traditional plan, the incentive plan will increase it; if the level was "too high," the incentive plan will trim preventive care to a socially efficient level.

Thus far, incentive-based health insurance plans have been used sparingly. Typically they have been introduced by self-insured employers seeking innovative ways to control health care costs, though the Mendocino Plan highlights the fact that only partial self-insurance may be needed to implement such an approach. The current limited federal experiment with MSAs could lead to wider support for incentive plans, but state-mandated coverage requirements and traditional insurers' lack of enthusiasm pose significant barriers. The results summarized above suggest that experiments with incentive-based health insurance could lead to beneficial changes in the structure of insurance, but success of these experiments will depend on contract features that are shaped by market conditions as well as federal and state restrictions. Further theoretical and empirical studies of how such parameters influence plan selection and utilization of services within plans could be useful in designing efficient incentive plans and refining policy restrictions.[19]

---

[19]19.   Van de Ven and Van Vliet (1995) empirically analyze the role of risk-adjusted premiums in reducing the potential for adverse selection in insurance markets. The plans that are analyzed contain no incentive component, but the authors' survey-based methods may be useful in establishing the parameters of an efficient incentive-based health plan.

References

1. Bogetic, Z., and D. Heffley, "Reforming Health Care: A Case for Stay-Well Health Insurance," Policy Research Working Papers, Country Operations, The World Bank, WPS 1181, September 1993.

2. Ehrlich, I., and G. Becker, "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy*, Vol. 80, No. 4 (1972), pp. 623-648.

3. Ellis, R., and R. McGuire, "Optimal Payment Systems for Health Services," *Journal of Health Economics*, Vol. 9, No. 4 (1990), pp. 375-396.

4. Hsiao, W., "Medical Savings Accounts: Lessons From Singapore," *Health Affairs*, Vol. 14, No. 2 (Summer 1995), pp. 260-266.

5. Kamien, M., and N. Schwartz, "Payment Plans and the Efficient Delivery of Health Care Services," *Journal of Risk and Insurance*, Vol. 40, No. 3 (September 1973), pp. 427-436.

6. Landes, W., and R. Posner, *The Economic Structure of Tort Law* (Cambridge, MA: Harvard University Press, 1987).

7. Massaro, T., and Y. Wong, "Positive Experience With Medical Savings Accounts in Singapore," *Health Affairs*, Vol. 14, No. 2 (Summer 1995), pp. 267-272.

8. Newhouse, J., "Policy Watch: Medicare," *Journal of Economic Perspectives*, Vol. 10, No. 3 (Summer 1996), pp. 159-167.

9. Nichols, L., "Medical Savings Accounts And Risk Segmentation," *Health Affairs*, Vol. 14, No. 2 (Summer 1995), pp. 275-276.

10. Ozanne, L., "How Will Medical Savings Accounts Affect Medical Spending?" *Inquiry*, Vol. 33, No. 3 (1996), pp. 225-236.

11. Pauly, M., and J. Goodman, "Tax Credits for Health Insurance and Medical Savings Accounts," *Health Affairs*, Vol. 14, No. 1 (Spring 1995), pp. 126-139.

12. Rothschild, M., and J. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Imperfect Information," *Quarterly Journal of Economics*, Vol. 90, No. 4 (1976), pp. 629-650.

13. Selden, T. "A Model of Capitation," *Journal of Health Economics*, Vol. 9, No. 4 (1990), pp. 397-409.

14. Swartz, K., "Medical Savings Accounts and Research," *Inquiry*, Vol. 33, No. 3 (1996), pp. 216-219.

15. The Urban Institute, "Winners and Losers Under Medical Savings Accounts," *Policy and Research Report*, Summer-Fall (1996), pp. 16-18.

16. Van de Ven, W., and R. van Vliet, "Consumer Information Surplus and Adverse Selection in Competitive Health Insurance Markets: An Empirical Study," *Journal of Health Economics*, Vol. 14, No. **2** (1995), pp. 149-169.

17. Zweifel, P., *Bonus Options in Health Insurance* (Boston, MA: Kluwer Academic, 1992).