Policy in competitive insurance markets: incentivizing risk sharing and cost efficiency

April 16, 2012

Ross Green

Professor Curtis R. Taylor, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University

Durham, North Carolina

Acknowledgements

I am grateful for the considerable contributions of two superior economists: my advisor Curt Taylor and my brother Brett Green. Without their insights, advice, encouragement and patience, this thesis would have remained incoherent or incomplete. All errors are strictly mine.

Abstract

In the setting of a population with heterogeneous risk of illness, informational asymmetries in a competitive health insurance market can cause the gains from risk sharing to fall short of social optimality in equilibrium. Traditional policies meant to address the under-provision of insurance, like mandating open enrollment or community-rated premiums, can be prohibitively costly or impossible to implement. I consider three policy regimes in the context of a competitive insurance industry in which firms maximize profits by exerting effort to monitor the provision of health care. When multiple risk types are present in the population, I find that a subsidy rule based on the marginal costs of insuring high risks can induce a Pareto-improvement to risk sharing gains, at a cost to the efficiency of health care provision. The novelty of the subsidy rule lies in the way it incentives pooling equilibria.

JEL classification: I0; I13; I18

Keywords: insurance contracts; private information; health care efficiency

1 Introduction

Markets for health insurance contracts have been a prime example of the importance of information since Rothschild and Stiglitz (1976) demonstrated the fragility of the equilibrium set. The large and increasing prices of health care in the United States, attributable to a number of factors including technological innovation, are at least in part the result of insurance contracts: specifically, the difficulties of resolving informational asymmetries in insurance markets, and of monitoring the application of contract terms to prevent overutilization of health care services. The increase in costs are not insignificant, and have demanded the attention of policymakers concerned with participation in insurance plans, as evidenced by the passage of the Patient Protection and Affordable Care Act in 2010. In this paper, I consider a competitive insurance market with private information and monitoring and management costs. I discuss the effects of different policy interventions on equilibrium quantities of insurance and incentives for managing costs of providing care. I find that a subsidy that reimburses insurers for the marginal cost of high types can induce a second-best outcome at lower cost than similarly-intentioned policies like open enrollment requirements.

2 Review of the literature

Newhouse, et al. (1997) identify empirical incentives for risk selection faced by health management organizations contracted by the U.S. government to enroll elderly Medicare beneficiaries. Their evidence illustrates the "equivalence principle" described by Van de ven, et al. (2000): in unregulated competitive markets, insurers will only extend insurance contracts such that economic profits are non-negative. In populations with variance in expected health care costs, premiums must be risk-rated for insurance to be extended universally. As Rothschild and Stiglitz (1976) first established, in the absence of information about risk types, insurers will attempt to cream-skim by distorting their benefit package or price-discriminate to sort out risk types; in this context, a separating equilibrium may or may not be stable, although under different assumptions, it can be shown that the separating equilibrium always exists (Duby and Geanakoplos, 2002). The same result occurs when regulation of the insurance industry mandates open enrollment to health plans or prohibits risk-rated premiums. Such regulation yields, both theoretically and empirically, sub-optimal participation in insurance plans. What's more, as noted by Newhouse (1982) and Pauly (1984), informational requirements likely make enforcement of open enrollment requirements impossible.

One alternative to open enrollment, designed to extend full insurance universally at a uniform price, focus on government transfer systems: in lieu of offering risk-rated premiums, insurers receive capitation payments, allocated based on risk-adjustment formulas that account for subscriber-level demographic information. The primary issue is that the observable characteristics on which formulas are based are often imperfect signals of individuals' health status. Frank, et al. (2000) introduce the concept of a shadow price of providing specific health care service—which accounts for expected capitation payments net of the market price of provision—through which managed care plans distort the benefit packages they offer. Frank, et al. (2000) use data from the Medicaid program in Michigan to estimate this shadow price, and show that risk-adjusted capitation fails to reduce selection incentives unless it equalizes the expected profits of provision over all health care services. Glazer and McGuire (2000) argue that, rather than providing payments that use regressions of certain demographic and diagnostic information to approximate patients' average expected cost, governments can reduce selection through incentives by prospectively "overpaying" insurers to take on high-risk subscribers.

On the one hand, a system of prospective capitation payments causes insurers to internalize transfer payments into their budget constraint, and thus incentivizes them to exert the optimal amount of effort in monitoring and managing provision of health care to their subscribers. But the practical limits of risk-adjustment formulas create an information problem for governments seeking to minimize risk-selection through capitation payments. Newhouse (1996a) notes that existing capitation formulas fail to explain the majority of between-person variance in expected health care costs. Health care providers have more information about potential subscribers than a capitation formula can incorporate, and hence insurers have incentives to "cream-skim" the lowest risk types and "dump" riskier types with higher expected costs. On the other, systems of cost-reimbursement remove insurers' incentive to improve the risk mix of their subscribers. Because insurers will not internalize any portion of reimbursed costs, however, they are not incentivized to exert the economically efficient amount of health care cost-reduction effort. Newhouse (1996a) described the dilemma facing reformers as the *selection-efficiency trade-off*.

Various forms of cost reimbursement, usually in the form of a common cost-sharing pool, to which insurers contribute some percentage of subscribers, have been explored in the literature as alternatives to improving ex ante risk adjustment. Van de Ven and van Vliet (1992) propose a cost-sharing plan whereby insurers designate a portion of high-risk subscribers for whom all costs will be reimbursed. Van Barneveld, et al. (2001) explore three forms of mandatory pooling as a cost-sharing mechanism: cost-sharing for high-risk subscribers, cost-sharing for subscribers with the highest demonstrated costs (excess-of-loss cost-sharing), and proportional cost-sharing. Van Barneveld, et al. (2001) find that risk-sharing for high-risks yields second-best outcomes on selection and efficiency using data from a Dutch sickness fund, though they do not account for the potential for distortion of the benefit package, or "stinting." Alternatively, Kifmann and Lorenz (2005) derive an optimal cost-reimbursement function that considers the incentives for efficiency and risk selection at every cost level. Using data from a Swiss health fund, Kifmann and Lorenz (2005) find that, rather than the outlier risk-sharing suggested by the literature, cost-reimbursement should be implemented only up to a threshold cost level. They attribute this result to the fact that outlier risk-sharing yields cost-sharing independent of the actual distribution of health care costs.

Sappington and Lewis (1999) consider a model where procurers of health care offer suppliers a menu of reimbursement options at the outset, a process they describe as "subjective risk adjustment," and find that both prospective payments and cost-sharing are components of the optimal transfer system. Ma (1994) first introduced a mixed payment scheme as a means of maximizing efficiency of provision of health care. Expanding on this result, Barros (2003) introduces a scheme in which fixed capitation payments are paid to insurers at the outset, and an ex post adjustment fund is used to compensate insurers for variations in their realized costs. In his model, the ex post fund induces both efficient provision of care and universal insurance, though this result rests on the assumption of homogeneity in risk type.

This paper follows these strands of literatue in examining the effects of various transfer systems on health insurance market outcomes. I evaluate three transfer systems in a similar context: a competitive insurance industry in which firms maximize profits through cost containment effort. I also allow for heterogeneity in risk type, requiring the optimal policy to artificially induce a pooling equilibrium.

3 The model with one type

The model describes the behavior of three agents in a competitive market for health insurance contracts: a person purchasing the contract, a firm selling the contract, and a government that regulates firms. People maximize their utility by purchasing a contract, according to the probability that they incur illness and their income. Firms offer a menu of contracts and maximize profit by selecting their cost containment effort. The government maximizes social utility by implementing policy. I define social utility as a simple sum of each person's utility.

The government moves first by setting its policy on the insurance industry. Firms move second, setting their effort level and offering a menu of contracts as a best response to the government's policy choice. Contracts are defined by the rate of insurance$x \in [0,1]$ and the premium p. People then purchase contracts from the menu and become ill with a given probability. Finally, firms fulfill the contract by providing medical care.

First, I present the model in a setting where the population exhibits one risk type: that is, everyone becomes ill with the same probability. In this setting, insurers offer only one type of contract. I show that competitive equilibrium corresponds to the social optimum, and thus that the government has no incentive to use policy to effect a different outcome.

Individual utility A person j incurs one type of illness with probabilityϕ_j. Utility is a function of the person's illness probability, his incomey_j, and his insurance contract. Utility in the model is *state-dependent,* because the marginal utility of income in the ill state depends on the rate of insurance. Expected utility is:

$$E(U_j) = \phi_j x (y_j - p) + (1 - \phi_j)(y_j - p).$$

With full insurance, j is indifferent toward becoming ill: at $x = 1$, he gets utility $y_j - p$ in both the ill and healthy state. People know their own probability of illness, and this information is private (e.g., j knows ϕ_j, but firm i does not).

Firm profits Firms make revenues from premiums and experience costs from providing medical care to subscribers. Firms also experience a sunk cost from the exertion of cost-containment effort $e \in [0,1]$. Think of e as a firm's effort to manage and monitor the provision of medical care to its subscribers. Exertion of effort lowers the firm's costs of providing care by preventing overutilization, either by providers seeking to induce demand (e.g., by recommending unnecessary MRIs) or by subscribers who do not fully internalize the costs of care (e.g., buying brand-name rather than generic). Effort level affects costs ambiguously: the sunk cost is increasing in effort, but costs of providing care are decreasing in effort. Firm effort is private information and cannot be observed by the government. In the setting with one risk type, firm i's profits are:

$$\Pi_i = p - x(0.5e_i^2 + (1 - e_i)\phi c).$$

The first term inside parentheses on the right corresponds to the sunk cost of effort. The second term describes the cost of providing care, which depends on subscribers' probability of illness and a cost parameter $c > 0$. Firms maximize profits according to the first-order condition for effort choice:

$$\frac{\partial \Pi_i}{\partial e_i} = x\left(\phi c - e_i\right) = 0 \Rightarrow e_i^\star = \phi c.$$

The second-order condition for a maximum is satisfied. As the expected costs of providing care rise, so too does the profit-maximizing level of effort. Firms adjust their effort to reflect their beliefs about the illness probabilities of their subscribers. Substituting the profit-maximizing effort level back into the profit function yields the expected costs per insurance contract, γ:

$$\gamma \equiv \phi c - 0.5(\phi c)^2.$$

I restrict expected costs per contract to be positive ($\gamma > 0$), which obtains if and only if:

$$1 - 0.5\phi c > 0.$$

In equilibrium, premiums are set according to the zero profit condition:

$$p^\star = \gamma x.$$

The insurance industry is competitive and all firms have the same cost function. In this setting, where every buyer of insurance has the same probability of illness, competition implies that all firms exert the same level of effort: $e_i^\star = e^\star, \forall i$.

Social Optimum The social optimum corresponds in this setting to the competitive equilibrium described above: j can purchase his desired rate of insurance at a premium determined by ϕ_j, while i fully internalizes its costs of providing care and maximizes profit with $e_i^\star$. In this setting, an unregulated market allows for open enrollment among insurance purchasers and properly incentivizes effort. Substituting the zero profit condition into the utility function yields the following maximization program for the government:

$$\max_x W_1 = \phi x \left(y - \gamma x\right) + \left(1-\phi\right)\left(y - \gamma x\right).$$

In the absence of asymmetric information, each person can individually maximize his or her own utility. The first-order condition $\frac{\partial W_1}{\partial x}$ yields j's optimal insurance purchase:

$$x_j^{OPT} = \frac{y_j}{2\gamma} - \frac{1-\phi_j}{2\phi_j}.$$

The purchase is unambiguously increasing in income y and decreasing in the expected costs of full insurance γ. Substituting for γ, it is apparent that j maximizes by purchasing at least some insurance $\left(x_j^{OPT} > 0\right)$ if his income is:

$$y_j > c(1 - 0.5\phi_j c)(1-\phi_j),$$

and j reaches the corner solution of full insurance $\left(x_j^{OPT} = 1\right)$ if his income is:

$$y_j > c(1 - 0.5\phi_j c)(1+\phi_j).$$

For the remainder of the paper, income is assumed to exceed this threshold level. As long as insurance contracts are priced based on j's illness probability (that is, no informational asymmetries exist), j maximizes utility by purchasing full insurance. When I consider more than one type in the subsequent sections, the inability of relatively low risk consumers to purchase full insurance motivates the government's intervention.

Substituting the zero profit condition back into the optimal insurance purchase yields $p^\star$ in terms of income, probability of illness and the cost parameter c:

$$p^\star = 0.5(y - (c - 0.5(\phi c)^2)(1-\phi).$$

Differentiating $p^\star$ with respect to ϕ gives

$$\frac{\partial p^\star}{\partial \phi} = 0.5c(1-0.5\phi c) + 0.25c^2(1-\phi) > 0.$$

The first term on the right is positive by the parameter restriction on γ, while the second term is positive by inspection. Higher probabilities of illness are unambiguously associated with higher insurance premiums. The indirect utility that j derives from purchasing his optimal contract $V(x^{\star}, p^{\star})$ is evidently greater than the utility of purchasing no insurance:

$$V(x^{\star}, p^{\star}) = \phi_j \gamma (x^{\star})^2 + y_j (1 - \phi_j) > y_j (1 - \phi_j).$$

I have established that if only one risk type is present, each firm exerts the same level of effort. In fact, there need not be full information of illness probabilities across the population to ensure this result. Less restrictively, if j rationally perceives his own probability of illness as ϕ such that $E(\phi_j) = \phi$ and earns income y_j such that $y_j > c(1 - 0.5\phi c)(1 + \phi)$, then $x_j^{OPT} = 1, \forall j$. All firms face the same provision costs in expectation; according to the timing of the model, their effort choice then converges on $e^{\star}$. A firm l that deviates by setting $e_l^{\star} < \phi c$ can break even on any given contract only by offering $p_l^{\star} > p^{\star}$. In a competitive industry, such a firm would not attract any subscribers.

In this setting, the market mechanism yields the social optimum without regulation. As suggested by much of the literature, if the population is homogeneous with respect to probability of illness, insurance markets do not exhibit problems of adverse selection and market unraveling—the so-called the insurance market "death spiral." Without any type of market failure, there is no need for government intervention in the market.

4 The model with heterogeneity over health risk

4.1 The model without regulation

I now turn to a more problematic setting: one where insurance buyers are heterogeneous in their probability of illness. In the model with two risk types, the population exhibits two illness probabilities (ϕ_H,ϕ_L) that occur with frequencies $(\Lambda, 1 - \Lambda)$. Since illness probability is private information, firms are now at an informational disadvantage and can determine their subscribers' type only by conducting first-degree price discrimination. Firms offer two insurance contracts, (p_H, x_H) and $(p_L, x_L < 1)$, such that high- and low-risk subscribers are perfectly sorted. The proportion λ_i of firm i's subscribers have illness probability ϕ_H, and $(1 - \lambda_i)$ have illness probability ϕ_L. It is assumed without loss of generality that firms have identical risk mixes: $\lambda_i = \Lambda, \forall i$. I demonstrate that competitive equilibrium does not correspond to a social optimum in the two type setting.

Insurer profits Firm i exerts a dimension of effort for each contract type they issue, in this case $e_{iH}$and e_{iL}. Firm's profit function remains otherwise the same in the two type setting. Firm i's profits are:

$$\Pi_i = \lambda_i \left[p_H - x_H \left(0.5e^2 + (1 - e_{iH})\, \phi_H c\right)\right] \\ + (1 - \lambda_i) \left[p_L - x_L \left(0.5e^2 + (1 - e_{iL})\, \phi_L c\right)\right].$$

The first order conditions governing the profit-maximizing choices of $e_H^\star$ and $e_L^\star$ yield the following effort levels:

$$e_H^\star = \phi_H c \qquad e_L^\star = \phi_L c.$$

To capture insurers' incentives for cost containment effort relative to the one-type case, it is instructive to compute a total effort level $e_T^\star$as a weighted average of $e_H^\star$ and $e_L^\star$. Define $\bar{\phi}_i$ as the average probability of illness of the insurer i's set of subscribers:

$$\bar{\phi}_i \equiv \lambda_i \phi_H + (1 - \lambda_i)\, \phi_L.$$

In equilibrium, total effort level is:

$$e_T^\star = \bar{\phi}_i c.$$

As shown in the one-type case, firm effort increases in the its subscribers' probabilities of illness, and by extension the costs of providing care. Insurance firms in the model are the sole suppliers of health care; people cannot, for example, forego insurance purchase health care on the spot market if they fall ill. This specification is made for clarity of exposition. But the model replicates reality in an important sense: insurance firms, who have a greater incentive to limit health care costs than providers or subscribers, bear the responsibility of managing costs. Firms whose subscribers have relatively higher probabilities of illness expect higher costs of providing care, and exert relatively great effort to mitigate costs. Their efforts at containing expected costs, then, should be expected to be relatively less significant, especially when effort implies a non-negligible fixed cost.

Substituting the type-specific profit-maximizing effort choices back into two-type profit function again yields the insurance-rate-adjusted expected costs for each contract type:

$$\gamma_H \equiv \phi_H c - 0.5(\phi_H c)^2, \qquad \gamma_L \equiv \phi_L c - 0.5(\phi_L c)^2.$$

Like in the one type case, in equilibrium, prices are set in accordance with zero profits to firms:

$$p_H^\star = \gamma_H x_H \qquad p_L^\star = \gamma_L x_L.$$

Social optimum In the one type setting, the definition of the social optimum as utility summed across the population proved to be compatible with the equilibrium in a competitive market. In this setting, information asymmetry drives a wedge between the social optimum and the realized equilibrium. Substituting in the zero profit and optimal effort conditions, the social optimum in the two type setting solves the maximization program:

$$\begin{aligned}\max_{x_L, x_H} W_2 = \Lambda & \left[\phi_H x_H \left(y_j - \gamma_H x_H\right) + \left(1-\phi_H\right)\left(y_j - \gamma_H x_H\right)\right] \\ & + \left(1-\Lambda\right)\left[\phi_L x_L \left(y_k - \gamma_L x_L\right) + \left(1-\phi_L\right)\left(y_k - \gamma_L x_L\right)\right].\end{aligned}$$

The first order conditions $\frac{\partial W_2}{\partial x_H} = 0$ and $\frac{\partial W_2}{\partial x_L} = 0$ yield (x_H, x_L) that solve W_2:

$$x_H^{OPT} = \frac{y_j}{2\gamma_H} - \frac{1-\phi_H}{2\phi_H} \qquad x_L^{OPT} = \frac{y_k}{2\gamma_L} - \frac{1-\phi_L}{2\phi_L}.$$

I return now to the assumption from the previous section that people maximize by purchasing full insurance if it is fairly priced, regardless of their type. For person j with illness probability $\phi_j = \phi_H$, this implies an income of y_j such that $x_H^{OPT} = 1$:

$$y_j \geq \frac{\gamma_H \left(1+\phi_H\right)}{\phi_H},$$

and, similarly, for person k with illness probability $\phi_k = \phi_L$, an income y_k such that $x_L^{OPT} = 1$:

$$y_k \geq \frac{\gamma_L \left(1+\phi_L\right)}{\phi_L}.$$

At the social optimum, both types purchase full insurance. Formally, the contracts traded under the social optimum are $\left(x_H^{OPT} = 1, p_H^{OPT} = \gamma_H\right)$, for the high risk person j, and $\left(x_L^{OPT} = 1, p_L^{OPT} = \gamma_L\right)$, for the low risk person k. Given this definition of the social optimum, I now define the two type equilibrium.

Competitive equilibrium In order to maximize producer surplus in the absence of information about risk types, firms must price-discriminate by offering quantities of insurance $x_L^\star$ and $x_H^\star$ determined by incentive compatibility constraints:

$$\phi_H x_H(y-p_H)+(1-\phi_H)(y-p_H) \geq \phi_H x_L(y-p_L)+(1-\phi_H)(y-p_L),$$

$$\phi_L x_H(y-p_H)+(1-\phi_L)(y-p_H) \leq \phi_L x_L(y-p_L)+(1-\phi_L)(y-p_L).$$

.

Given $\phi_H > \phi_L$, these constraints obviously preclude an equilibrium in which $x_L^\star = 1$. At full insurance, utility is the same in both the healthy and ill states of nature, so the left sides of each incentive compatibility constraints are equal. Person k with illness probability ϕ_L is demonstrably better off with any partial insurance plan $x_L < 1$ than is person j with illness probability ϕ_H. At the equilibrium value of x_L, the incentive compatibility constraint for high risk types holds with equality, and the constraint for low risk types does not bind:

$$y-p_H = \phi_{pH} x_L(y-p_L)+(1-\phi_H)(y-p_L) < \phi_L x_L(y-p_L)+(1-\phi_L)(y-p_L).$$

Substituting into the binding incentive compatibility constraint gives

$$x_L^\star = \frac{\phi_H\left(y-p_L\right)+p_L-p_H}{\phi_H\left(y-p_L\right)} = 1+\frac{p_L-p_H}{\phi_H\left(y-p_L\right)} < 1,$$

and, by the restriction on y_k, $x_L^\star < x_L^{OPT}$. Indirect utility is unambiguously higher for low-risk types under x_L^{OPT} than $x_L^\star$ (Appendix A contains an algebraic proof of this result):

$$V\left(x_L^{OPT}, p_L^{OPT}\right) = \phi_L\left(y_k-\gamma_L\right)+\left(1-\phi_L\right)\left(y_k-\gamma_L\right) = y_k-\gamma_L.$$

$$V\left(x_L^\star, p_L^\star\right) = \left(y_k-\gamma_L\right)+\gamma_L\left(\frac{p_L-p_H}{\phi_H\left(y_k-p_L\right)}\right).$$

Looking at the indirect utility of $V\left(x_L^\star, p_L^\star\right)$, the second term on the right is negative by inspection. The insurance market equilibrium $\left[\left(x_L^\star, p_L^\star\right), \left(x_H^\star, p_H^*\right)\right]$ does not maximize the social welfare function W_2. This result establishes the problem from the perspective of the government: when two types are present in the population, equilibrium does not correspond to a social optimum. High risk types can purchase the optimal quantity of insurance, but low risk types experience utility losses due to insurer price discrimination. I have shown in the context of state-dependent utility, since the marginal utility of income is increasing in the quantity of insurance purchased, but the result holds for any strictly concave utility function as well. As first discussed by Rothschild and Stiglitz (1977) in their landmark paper on imperfect information, the presence of high risk types with private information imposes a purely dissipative negative externality on low risk types. Equilibrium in the insurance market is demonstrably Pareto-inferior.

4.2 Reaching the social optimum through regulation

I have defined the socially optimal outcome for a competitive health insurance industry facing a sunk cost from effort, and shown that equilibrium corresponds to the social optimum only when the population is homogenous in risk type. Thus far, the government has yet to play a role in the story. In this section, I consider whether government policy can effect a social optimum, or, failing this, a Pareto improvement over the equilibrium that obtains in the two type setting.

In order to focus the analysis on practically feasible policies, I consider only subsidies a) to firms, rather than insurance buyers, and b) with linear payment schedules. The government cannot observe illness probabilities, nor can it observe or directly regulate firms' effort level; the government can, however, observe the population parameter Λ, the terms of all traded insurance contracts (x_n, p_n), and the costs to firms of providing care. In addition to linear subsidies, the government can impose an open enrollment rule—under which firms cannot reject anyone at a given premium—and restrict the number of different contracts firms can offer. Social welfare maximization is constrained by competitive equilibrium conditions and by firms' optimal effort condition, and is given generally by:

$$\begin{aligned}\max_{x_L, x_H} W_2 = {} & \Lambda \left[\phi_H x_H \left(y_j - p_H^\star\right) + \left(1-\phi_H\right)\left(y_j - p_H^\star\right)\right] \\ & + \left(1-\Lambda\right)\left[\phi_L x_L \left(y_k - p_L^\star\right) + \left(1-\phi_L\right)\left(y_k - p_L^\star\right)\right].\end{aligned}$$

I consider now three subsidy rules, and determine whether implementation of any of the three by the government in the first stage yields the social optimum or a Pareto-superior second-best outcome.

4.2.1 Mandatory pooling with proportional cost reimbursement

The policy combines the restriction that firms offer only one type of contract with a proportional cost reimbursement subsidy. In response to the restriction on contracts, firms offer one contract,$(\bar{x} = 1, \bar{p})$, such that

$$\Pi_i = \bar{p} - \left[0.5e_i^2 + (1 - e_i)\,\bar{\phi}c\right].$$

The restriction prevents firms from using contracts to identify risk types. The uniform premium $\bar{p}$ implies that firms expect positive profits from low risk subscribers and losses from high risk subscribers. The result is a forced pooling equilibrium where participation of low risk types in the market subsidizes high risk types. According to the timing of the model, a cost reimbursement subsidy is necessarily a function of the effort level firms select after the policy is announced, and is given by:

$$S_1(e_i) = \alpha\,(1 - e_i)\,\bar{\phi}c,$$

where the government reimburses some proportion $0 \leq \alpha \leq 1$ of a firm's expected costs of provision $(1 - e_i)\,\bar{\phi}c$. According to the timing of the model, the government must announce the rule prior to firms selecting their profit-maximizing effort level. Profits under this policy are:

$$\Pi_i = \bar{p} - \left[0.5e_i^2 + (1 - \alpha)(1 - e_i)\,\bar{\phi}c\right],$$

and the first-order condition for effort choice yields the profit-maximizing effort level under S_1:

$$e^{\star}_{i|S_1} = (1 - \alpha)\,\bar{\phi}c.$$

The shortcomings of this policy are clear. Any nonzero value of α reduces firms' incentive to exert effort and prevents efficient provision of care. Distorting incentives this way creates an efficiency loss that can be calculated by substituting $e^{\star}_{i|S_1}$ into the total expected costs per contract:

$$\bar{\gamma}(S) = S_1 + 0.5\left(e^{\star}_{i|S_1}\right)^2 + (1 - \alpha)\left(1 - e^{\star}_{i|S_1}\right)\bar{\phi}c = \bar{\phi}c - 0.5\left(\bar{\phi}c\right)^2 + 0.5\left(\alpha\bar{\phi}c\right)^2.$$

The efficiency cost of the subsidy is given by the quantity $\left(\alpha\bar{\phi}c\right)^2$, which is positive for any nonzero α. The efficiency cost is increasing and concave in α—differentiating expected costs with respect to α yields:

$$\frac{\partial \bar{\gamma}}{\partial \alpha} = 2\alpha \left(\bar{\phi}c\right)^2 \geq 0.$$

This is only part of the problem with Policy 1. The restriction to one contract means that premiums cannot be risk-adjusted; depending on the disparity between ϕ_L and ϕ_H, low risk types may opt out of the insurance market altogether. If $\alpha = 0$, low risk types purchase $(\bar{x} = 1, \bar{p})$ if and only if:

$$\phi_L \bar{x}(y_k - \bar{p}) + (1 - \phi_L)(y_k - \bar{p}) \leq (1 - \phi_L)y_k.$$

Low risk types opt out of the insurance market completely whenever this incentive compatibility constraint does not hold. Even when it does hold, low risk types could still experience a utility loss from being forced to pool with high types, if:

$$\phi_L \bar{x}(y_k - \bar{p}) + (1 - \phi_L)(y_k - \bar{p}) \leq \phi_L x_L(y_k - p_L) + (1 - \phi_L)(y_k - p_L).$$

In this case, the policy does not represent a Pareto-improvement over unregulated equilibrium. High risk types are made better off at the expense of low risk types.

A nonzero α doesn't do much to sweeten the deal for low risk types. In equilibrium, firms charge a premium $\bar{p}$ equal to the costs of insurance that firms internalize. Substituting for $e_i^\star$, the zero-profit condition for $\bar{p}$ is:

$$\bar{p} = (1 - \alpha)\,\bar{\phi}c\,[1 - 0.5\phi c\,(1 - \alpha)]\,.$$

Differentiating $\bar{p}$ with respect to α yields:

$$\frac{\partial \bar{p}}{\partial \alpha} = \bar{\phi}c\left[(1 - 2\alpha)\,\bar{\phi}c - 1\right].$$

Recall the restriction from the first section that insurance contracts have positive expected costs: $1 - 0.5\phi c > 0$. For $\phi c = \bar{\phi}c$, the restriction implies that $\bar{p}$ is unambiguously decreasing only if $\alpha \geq 0.25$. That is, given that firms can offer only one insurance contract, the government must subsidize fully one-quarter of the total costs of the contract in order to reduce the equilibrium premium with certainty. Up to $\alpha = 0.25$, the indirect effect of the subsidy rule on firms' effort level is greater than the direct effect of the payment itself.

This finding is intuitive and illustrates the importance of firms internalizing costs when selecting their profit-maximizing effort level. For any $\alpha > 0$, firms no longer fully internalize the costs of provision and the insurance industry experiences an efficiency cost. Similar conclusions regarding proportional cost reimbursement can be found in the literature, including in Barros (2003).

4.2.2 Mandatory pooling with reimbursement of cost differential

Given the sub-optimality of proportional cost reimbursement, I now explore an alternative subsidy. Again, the government mandates pooling, so that firms can offer at most one type of insurance contract. This policy differs from the prior one in the form of the subsidy. Rather than reimbursing a proportion of total costs, the government under this policy reimburses the difference in expected costs between insuring high risk and low risk types. Define the subsidy under this policy as:

$$S_2\left(e_i\right) = \Lambda c \left(1 - e_i\right)\left(\phi_H - \phi_L\right).$$

One consequence of mandatory pooling is that it prevents firms from risk-adjusting premiums. When cost reimbursement is a simple proportion of total costs, mandatory pooling makes low risk types worse off because their premium exceeds the actuarially fair level, with the surplus going to subsidize high risk types. S_2 alleviates those equity concerns by reimbursing firms for the marginal cost of insuring high risk types. Firm i underS_2 has profits:

$$\begin{aligned}\Pi_i = \bar{p} - \Big[0.5e_i^2 + \lambda\left(1 - e_i\right)\left(\phi_H - \phi_H + \phi_L\right)c + \left(1 - \lambda\right)\left(1 - e_i\right)\phi_L c\Big] \\ = \bar{p} - \Big[0.5e_i^2 + \left(1 - e_i\right)\phi_L c\Big].\end{aligned}$$

Recall here one of the assumption from Section 4 that $\lambda_i = \Lambda, \forall i$: essentially, that each firm's risk mix is identical to that of the population. Without regulation, firms price-discriminate to separate high and low risk types and risk-adjust premiums so that expected profits from both contracts are zero. The assumption that all firms have an identical risk mix is in reality unlikely to hold—some firms will have contract with a higher proportion of high risk types than others—but in the absence of regulation, allowing for variation in firms' risk mix does not fundamentally alter equilibrium. Each firm i still offers (x_L, p_L) and exerts $e^\star = \phi_L c$ for low risk types, and offers (x_H, p_H) and exerts $e_i^\star = \phi_H c$ for high risk types; no individual is made better or worse off, so social welfare is unaffected.

Under S_2, firms' risk mix has a real effect on profits, since the returns to insuring high and low risk types are no longer equal. Since the government cannot observe risk type, it must use the population parameter Λ to estimate firm i's proportion of high risk subscribers λ_i. When information about risk type is private—as has been assumed throughout this paper—firms are unable to cream-skim. Variation between population parameter and firm risk mix should be zero in expectation and randomly distributed:

$$E\left(\Lambda-\lambda_i\right)=0, \forall i.$$

The literature shows that even when firms are unable to cream-skim (under open-enrollment rules, for instance) they find other ways to select risks, like "stinting" benefit packages to make them less attractive to relatively risky consumers by excluding especially high-cost services. Under S_2, any form of risk selection that does not affect a firm's competitiveness in the marketplace would be profitable. Firms with the ability to screen for risk type maximize profits by insuring only those with illness probability ϕ_L. For now, I continue with the assumption that risk selection is impossible or prohibitively costly; in the concluding section, I discuss the implications of risk selection for the government's policy choice.

Returning to the profit function under S_2 , the first-order condition for effort gives the profit-maximizing effort level:

$$e^{\star}_{i|S_2}=\phi_L c.$$

Like the previous subsidy, S_2 reduces incentives to exert effort. With mandatory pooling and no subsidy, $e^{\star}_i=\bar{\phi}c>\phi_L c$. Again, the efficiency loss from the subsidy can be calculated by substituting $e^{\star}_{i|S_2}$ into the total expected costs per contract:

$$\bar{\gamma}_{S_2}=S_2+0.5\left(e^{\star}_{i|S_2}\right)^2+\left(1-e^{\star}_{i|S_2}\right)\bar{\phi}c=\bar{\phi}c-0.5\left(\phi_L c\right)^2+\Lambda\phi_L c^2\left(\phi_L-\phi_H\right).$$

The total expected costs per contract without a subsidy in place are:

$$\bar{\gamma}=\bar{\phi}c-0.5\left(\bar{\phi}c\right)^2+\Lambda\phi_L c^2\left(\phi_L-\phi_H\right)-0.5\Lambda^2c^2\left(\phi_H^2-\phi_L^2\right).$$

The fourth term on the right side of the equation, $0.5\Lambda^2c^2\left(\phi_H^2-\phi_L^2\right)$, is unambiguously positive and represents the marginal efficiency loss from the subsidy.

Despite the loss in the efficiency of care, the policy does yield a Pareto-improvement for both high and low risk types. Person j with illness probability ϕ_H enjoys full insurance, as in the unregulated equilibrium, but his utility is greater because his premium is relatively less expensive:

$$p_H=\phi_H c-0.5\left(\phi_H c\right)^2>\phi_L c-0.5\left(\phi_L c\right)^2=\bar{p}.$$

Meanwhile, person k with illness probability ϕ_L continues to pay a premium that is properly adjusted to his risk of illness, as in unregulated equilibrium, but can now purchase a full insurance contract—the same contract described as $\left(x_L^{OPT},p_L^{OPT}\right)$ in Section 4.

4.2.3 Reimbursement of cost differential with incentives for full insurance

Both of the policies discussed to this point included a constraint on firms in the form of mandatory pooling. Here, I consider an alternative mechanism to induce full insurance: making the subsidy payment contingent on the firms' offered rates of insurance. As before, the subsidy represents the expected difference between the costs of providing care to high risk types and the costs of providing it to low risk types. Define the subsidy as:

$$S_3 = \Lambda \frac{x_L}{x_H} c \left[(1 - e_H) \phi_H - (1 - e_L) \phi_L \right].$$

In order to collect the full cost differential, firms must set $x_L = x_H$. A firm attempting to separate risk types would legally be free to do so, but the subsidy disincentives this strategy: firms that offer full insurance universally receive the full subsidy payment and internalize only the costs associated with insuring low risk types, anyway. With the subsidy in place and $x_H = 1$, firm i's profits are:

$$\Pi_i = \lambda_i \left[p_H - 0.5e_H^2 - (1 - e_{iH}) \phi_H c + (1 - e_{iH}) \phi_H x_L c - (1 - e_{iL}) \phi_L x_L c \right] + (1 - \lambda_i) \left[p_L - x_L \left(0.5e^2 + (1 - e_{iL}) \phi_L c \right) \right]$$

Substituting the zero-profit condition for the low-risk type premium p_L and differentiating profits with respect to x_L, we get:

$$\frac{\partial \Pi}{\partial x_L} = \lambda c \left[(1 - e_H) \phi_H - (1 - e_L) \phi_L \right] + (1 - \lambda) \left(\gamma_L - 0.5e^2 - (1 - e_L) \phi_L c \right) > 0.$$

The second term on the right side of the equation is zero by definition of γ_L. The first term on the right side is positive by inspection. Firms maximize profits by setting $x_L = 1$ and offering a single full insurance contract. Costs associated with high-risk types are, as under the previous policy, the same in expectation as those associated with low-risk types. Once firms voluntarily pool their subscribers, they again exert effort in only one dimension. Profits can be functionally described by:

$$\Pi_i = \bar{p} - 0.5e_i^2 - (1 - e_i)\phi_L c.$$

Like the previous policy, the profit-maximizing effort choice is given byϕ_{LC}; the marginal loss of efficiency is the same under S_3 as it was under S_2. Given the assumption that both high and low risk types, faced with properly risk-adjusted premiums, maximize utility by purchasing full insurance, S_2 and S_3 induce the same outcome.

However, if some of the assumptions made to this point are relaxed, the properties of S_3 make it the most practical and flexible policy option of the three outlined above. If, for example, the government experiences costs from monitoring firms for compliance with the full insurance regulation, S_3 is strictly preferable to S_2. If the assumption that person j with illness probability ϕ_H maximizes by purchasing full insurance does not hold, such that

$$y_j < \frac{\gamma_H \left(1 + \phi_H\right)}{\phi_H},$$

high risk types instead maximize utility in unregulated equilibrium by purchasing some quantity of insurance $x_H < 1$. In this context, a rule that requires all contracts offer full insurance, as implied by S_2, is not social welfare maximizing. In order to mandate pooling at the utility maximizing level x_H^{OPT}, the government would need to be able to reliably monitor the utility maximizing insurance purchase for high-risk individuals—a purchase that may be sensitive to relative changes in health care prices over time. S_3 is less information-intensive; it shifts the burden of determining x_H^{OPT}onto firms, who converge to offer this value in equilibrium. Whereas S_2 necessarily yields a Pareto-improvement over unregulated equilibrium when $x_L^{OPT} = x_H^{OPT} = 1$, S_3 yields a Pareto-improvement more generally whenever $x_L^{OPT} = x_H^{OPT}$:

$$\frac{\gamma_H \left(1 + \phi_H\right)}{\phi_H} = \frac{\gamma_L \left(1 + \phi_L\right)}{\phi_L}.$$

The desirable properties of S_3 are thus demonstrated: it yields a Pareto-improvement in social welfare in any state where the demand for risk-adjusted insurance is the same across both risk types.

5 Concluding remarks

The market for health insurance contracts is fraught with the potential for myriad market failures. Private information limits gains from sharing health risks; demand for health care services is distorted by principal-agent problems and morally hazardous spending incentives; adverse selection prevents the highest risks from entering the market, while monitoring and loading costs make health insurance a prohibitively expensive proposition for individual buyers. Even beyond issues of efficiency, the pro-social policymaker must consider fundamental equity concerns, as well. Are individuals responsible for their own health condition, and to what extent can premiums be adjusted for risk, if at all? Is possession of health insurance a universal right or a tradable good?

This paper explores in depth one of these tradeoffs: maximizing the gains from risk-sharing versus minimizing the cost of health care provision. I model a competitive insurance industry supplying a population with heterogeneity over health risks, in which firms maximize expected profits in a given period by exerting cost-containment effort. In the interest of tractability, I specify consumers' utility as state-dependent; their marginal utility of income in the ill state is increasing in their rate of insurance. In the absence of any regulation, the existence of a sunk cost of effort does not resolve the dilemma of private information. The conditions for a separating equilibrium in which high risks are fully insured and low risks are partially ensured are established, and I derive a closed-form solution for the equilibrium rate of partial insurance. These basic results are robust to alternative specifications in which utility is strictly concave. I then consider the effects of three policies on a) firms' incentive to contain costs and b) the utility gains to consumers from sharing risks.

5.1 Results and policy implications

The first policy, S_1, under consideration combines two elements: mandating full insurance pooling and proportional reimbursement of total costs of provision. Consistent with the literature, I find that proportional cost reimbursement adversely affects firms' cost-containment effort, leading to inefficient provision of care. Such a policy is unlikely to benefit low risk types, who pay actuarially unfair premiums that subsidize high risk types in the pooling equilibrium. Finally, because this policy creates variation in the expected profits per contract, it violates the "equivalency principle" by creating incentives for adverse selection, though this possibility is not considered explicitly in the model.

The second policy, S_2, retains the full insurance mandate and replaces proportional cost reimbursement with a subsidy that could be termed "marginal cost reimbursement:" the government reimburses insurers for the marginal cost of insuring high risks relative to low risks. Such a policy allows both types to purchase full insurance at premiums that are actuarially fair or better, and so represents a Pareto-improvement over unregulated equilibrium. This policy also adversely affects insurer effort, and also creates incentives to risk selection where screening by firms is possible.

The third policy, S_3, also involves a marginal cost reimbursement subsidy, but achieves full insurance through incentives rather than through mandate. The novelty of S_3 lies in the conditionality of the subsidy. The marginal cost reimbursement is multiplied by the ratio $\frac{x_L}{x_H}$: the rate of insurance offered to low risks divided by the rate offered to high risks. In order to obtain full marginal cost reimbursement, insurers must offer a pooling plan. When both types desire full insurance, this policy yields the same results for efficiency and risk sharing as does S_2. However, if the utility-maximizing insurance contract stipulates anything less than full insurance, S_3 is relatively less information-intensive and thus strictly preferrable to S_2. S_3 becomes even more attractive relative to the alternatives if the government incurs costs from enforcing the full insurance mandate.

However, like the alternatives, S_3 violates the equivalency principle and thus incentives for adverse selection. The model assumes that screening by insurers is impossible or prohibitively expensive, and so abstracts away from the possibility of selection, though this is unlikely to be the case in real-world settings. The most intriguing potential for the application of S_3, then, lies not in the setting where information about risk type is strictly private, but where insurers are unable to use observable information to risk-adjust premiums. Equity-minded governments may hold that insurers cannot discriminate on the basis of race, gender, or place of birth, even though this information is easily observable. For instance, the Patient Protection and Affordable Care Act of 2010 stipulates a form of community-rated premiums: insurers are not allowed to increase premiums for consumers with pre-existing conditions. While such regulations might preclude insurers from obvious forms of adverse selection, like cream-skimming or risk-rating premiums, certain forms of risk selection—like manipulating benefit packages to discourage high risks from subscribing—may be difficult for a policymaker to monitor. S_3 represents a desirable alternative: by incentivizing insurers to pool risk types, it alleviates the costs to a government of enforcing a no-selection mandate. S_3 also benefits both the relatively risky and the relatively riskless consumers in comparison to mandatory pooling.

5.2 Avenues for future research

While S_3 has a number of desirable properties, it does decrease incentives for cost containment effort and lead to inefficient provision of care. Variation between the highest and lowest risks exacerbates this inefficiency. The challenge for researchers and policymakers is threefold: maintain the incentives for effort provided by competition among firms to ensure efficient provision of care; equalize the expected profits associated with different risk types to ensure the stability of a pooling equilibrium; and incentive higher rates of insurance in order to maximize the gains from risk sharing. Future research should continue to evaluate policy alternatives in search of a first-best rule, both in the setting described here and alternative settings—where, for example, the population exhibits n risk types, or health insurance purchasers face multiple health risks.

Appendix

Algebraic proof of $V\left(x_L^{\star}, p_L^{\star}\right) < V\left(x_L^{OPT}, p_L^{OPT}\right)$:

$$V\left(x_L^{\star}, p_L^{\star}\right) = \phi_L x_L \left(y_k - \gamma_L\right) + \left(1 - \phi_L\right)\left(y_k - \gamma_L\right)$$

Substituting for $x^{\star}$ from the incentive compatibility constraint:

$$V\left(x_L^{\star}, p_L^{\star}\right) = \phi_L \left(1 + \frac{p_L - p_H}{\phi_H \left(y - p_L\right)}\right)\left(y_k - \gamma_L\right) + \left(1 - \phi_L\right)\left(y_k - \gamma_L\right)$$

Rearranging terms:

$$V\left(x_L^{\star}, p_L^{\star}\right) = y_k - \gamma_L + \phi_L \left(\frac{p_L - p_H}{\phi_H \left(y - p_L\right)}\right)\left(y_k - \gamma_L\right)$$

Then substituting for y_k:

$$V\left(x_L^{\star}, p_L^{\star}\right) = y_k - \gamma_L + \phi_L \left(\frac{p_L - p_H}{\phi_H \left(y - p_L\right)}\right)\left(\frac{\gamma_L \left(1 + \phi_L\right)}{\phi_L} - \frac{\gamma_L \phi_L}{\phi_L}\right)$$

$$V\left(x_L^{\star}, p_L^{\star}\right) = y_k - \gamma_L + \left(\frac{p_L - p_H}{\phi_H \left(y - p_L\right)}\right)\gamma_L.$$

Since the third term on the right is negative by inspection, it is evident that this is lower than the utility associated with the optimal plan:

$$V\left(x_L^{OPT}, p_L^{OPT}\right) = \phi_L \left(y_k - \gamma_L\right) + \left(1 - \phi_L\right)\left(y_k - \gamma_L\right) = y_k - \gamma_L$$

References

[1] Barros, P.P. 2003. Cream-skimming, incentives for efficiency and payment system. *Journal of Health Economics* 22. 419-443.

[2] Frank, Richard G., Jacob Glazer, and Thomas G. McGuire. 2000. Measuring adverse selection in managed health care. *Journal of Health Economics* 19. 829–854.

[3] Glazer, Jacob, and Thomas G. McGuire. 2011. Gold and silver health plans: accommodating demand heterogeneity in managed competition. *Journal of Health Economics* 30. 1011-1019.

[4] Glazer, Jacob, and Thomas G. McGuire, 2000. Optimal risk adjustment in markets with adverse selection: an application to managed care. *American Economic Review* 90. 1055–1071.

[5] Kifmann, Mathias, and Normann Lorenz. 2011. Optimal cost reimbursement of health insurers to reduce risk selection. *Health Economics* 20. 532-552.

[6] Lewis, Tracy R. and David E.M. Sappington. 1999. Access pricing with unregulated downstream competition. *Information Economics and Policy* 11. 73-100.

[7] Newhouse, Joseph. 1982. Is competition the answer? *Journal of Health Economics* 1. 110-116.

[8] Newhouse Joseph. 1996. Reimbursing health plans and health providers: efficiency in production versus selection. *Journal of Economic Literature* 34. 1236–1263.

[9] Newhouse, Joseph, Meredith Beeuwkes Buntin, and John D. Chapman. 1997. Risk adjustment and medicare: taking a closer look. *Health Affairs* 16. 26–43.

[10] Pauly, Mark V. 1984. Is cream-skimming a problem for the competitive medical market? *Journal of Health Economics* 3. 87-95.

[11] Rothschild, Michael, and Joseph Stiglitz. 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics* 90. 629-649.

[12] Van Barneveld, Erik M., Rene C.J.A. van Vliet and Wynand P.M.M van de Ven. 2001. Risk sharing between competitive health plans and sponsors. *Health Affairs* 20. 253-262.

[13] Van de ven, Wynand P.M.M., Rene C.J.A. van Vliet, Frederik T. Schut, and Erik M. van Barneveld. 2000. Access to coverage for high-risks in a competitive insurance market: via premium restrictions or risk-adjusted premium subsidies? *Journal of Health Economics* 19. 311-339.

[14] van de ven, Wynand P.M.M. and Rene C.J.A. van Vliet. 1992. How can we prevent cream skimming in a competitive health insurance market? In *Health Economics Worldwide*. Eds. Zweifel P., Frech III H. Kluwer: Dordrecht. 23–46.