

Optimal risk adjustment with adverse selection and spatial competition

William Jack*

Department of Economics, Georgetown University, Washington DC 20057, United States

Received 1 October 2004; accepted 2 January 2006

Available online 19 May 2006

Abstract

Paying insurers risk-adjusted prices for covering different individuals can correct selection incentives and induce the market to provide optimal insurance policies. To calculate the optimal risk-adjusted prices we need to know (a) what the optimal policies are; (b) how much they cost; and (c) how competitive the market is. We examine these issues in a model with spatial heterogeneity and adverse selection. Market equilibrium is characterized, and delivery of the socially optimal insurance policies is possible, as long as providers are paid risk-adjusted fees for each individual they serve. When the payment can be made on the basis of an individual's risk, it should be sufficient to cover the expected cost of the socially optimal policy for that person, plus a mark-up. If payments can be made only on the basis of a partially informative signal, the optimal risk-based payments should be adjusted according to a simple linear transformation, identified by Glazer and McGuire [Glazer, J., McGuire, T., 2000. Optimal risk adjustment of health insurance premiums: an application to managed care. *American Economic Review* 90 (4), 1055–1071].
© 2006 Elsevier B.V. All rights reserved.

JEL classification: I11; I18

Keywords: Risk adjustment; Adverse selection; Spatial competition

1. Introduction

In the hope of providing stronger efficiency incentives and ensuring greater responsiveness to demand, many public services have recently been contracted out to private providers. For example, by July 2005, 12% of Medicare enrollees in the US obtained health insurance coverage

* Tel.: +1 202 687 0773.

E-mail address: wgj@georgetown.edu.

through private sector Health Maintenance Organizations (HMOs),¹ while a number of states have implemented school voucher schemes and other means of increasing school choice. In these kinds of markets, private providers are paid by an intermediary purchaser (the government) for each individual they serve. Alternatively, individuals may be allocated a voucher with a specified face value, that can be used to purchase services offered by a provider. A problem that many health economists have identified, and in recent years begun to model, arises when individuals are heterogeneous. If a uniform price is paid to providers on behalf of individuals they serve, or if vouchers have a uniform face value, incentives may arise for providers to restrict access of certain (high cost) individuals to services. If providers are required to offer a single standardized bundle of services, then non-discrimination and open enrollment regulations might mitigate against such actions.² On the other hand, individual heterogeneity typically means provision of a single bundle of services for all types of individuals is inefficient. Individuals at risk of heart disease need cardiovascular surgeons in their health plan, and parents of children with learning disabilities want special education services (Biglaiser and Ma, 2003). There is then a concern that if providers are afforded the flexibility of offering different bundles of services, they might strategically alter the mix not so much to efficiently match consumer preferences, but to induce unprofitable individuals to seek services elsewhere, while attracting more profitable clients.

Such selection-induced marketing policies could result in costly distortions to equilibrium service bundles. One response is for the purchaser to alter the prices paid on behalf of different consumers, in the hope of reducing the variation in profitability across individuals. This is referred to as risk adjustment in the health economics literature (e.g., Newhouse, 1996). Much of the risk adjustment literature has been empirical in nature, searching for good predictors of individual health care spending.³ But if different individuals would optimally consume different bundles of services, then predicting the cost of providing a given bundle to individuals with different characteristics is not necessarily enough information for the purchaser to optimally set prices.

This paper contributes to a growing literature on “optimal” risk adjustment mechanisms. We approach the question of how to optimally differentiate prices paid to providers on behalf of heterogeneous consumers in four steps. First, we ask what bundles of goods a social planner would directly allocate to the different consumers in maximizing social welfare. This optimization is performed under the incentive compatibility constraints imposed by the requirement that consumers can freely choose between the bundles that are provided. The second step is to model the equilibrium behavior of providers when they are paid a fixed uniform price for all individuals. In the recent literature on managed care behavior (Frank et al., 2000, hereafter FGM, Glazer and McGuire, 2002, hereafter GM02) providers are assumed to allocate resources by choosing good-specific shadow prices when competing. We show that this neither is consistent with profit maximization (Ma, 2003 also makes this point), nor does it necessarily respect incentive compatibility constraints imposed by open enrollment rules.⁴

¹ <http://www.kff.org/medicare/upload/Medicare-Advantage-April-20050Fact-Sheet.pdf>, although this percentage is expected to risk in 2006 due to changes in reimbursement rules. Murgolo (2002), reported that the proportion was 17% in 2000, and that in fact, conditional on having (geographic) access to an HMO, 25% of eligible individuals got insurance coverage through such an organization at that time.

² Shen and Ellis (2002) present a model in which health plans can perfectly cream skim—that is, they can identify costly patients and deny them coverage.

³ See Newhouse (2002, chapter 6) and Van de Ven and Ellis (2000) for a review of this literature.

⁴ Of course, some providers are non-profit, but this in itself should not imply that they offer allocatively inefficient plans.

Third, assuming the purchaser can perfectly identify an individual's relevant characteristics (his or her "type"), then given the equilibrium behavior of providers, prices can be differentiated by individual type to maximize the purchaser's objective function. The optimal prices might be called "optimal type-based prices." Finally, the fourth step is to recognize that the purchaser will generically have incomplete information on individual characteristics, and to address the statistical issue of how payments should be related to that incomplete information. The statistical correction of Glazer and McGuire (2000), hereafter GM00, is used for this purpose.⁵

In the model in the current paper individuals differ on two dimensions: first with respect to their risk of suffering bad health, and second on a continuous spatial dimension. There is no *ex ante* or *ex post* moral hazard (i.e., an individual's *ex ante* risk of bad health is exogenous, and her *ex post* health status is observable), and a representative insurance policy is defined by the quality of health care services provided under the policy in alternative health states. Consistent with the recent literature, health care is multidimensional, and insurance policies can promise high quality of some kinds of services (e.g., pediatrics), but low quality of others (e.g., mental health). The idea is that the quality of, say pediatric care, offered under an insurance policy is determined by the staffing and management policies of the insurer, and this level of quality is defined in the insurance package. Alternative packages might restrict access to certain physicians or ancillary services, thereby offering a potentially different quality of service.

Later, when we consider risk-adjusted payments, we will assume that providers cannot *directly* differentiate between (i.e., select) individuals on the basis of their observable characteristics, even if a purchaser conditions reimbursement rates on such characteristics. For example, we will allow a purchaser to reimburse a provider on the basis of the age and gender of an individual, but we will assume that providers cannot (i) deny enrollment to an individual based on age and gender; or (ii) explicitly differentiate the quality of services provided to individuals on the basis of their age and gender. This does not mean that an insurer cannot *indirectly* (in equilibrium), differentiate between (i.e., select) individuals according to their age/gender characteristics: a policy with good maternity and pediatric care but poor geriatric care would be attractive to young adults, but not to the aged. We simply assume that a provider cannot explicitly preclude the aged from choosing such a policy. Similarly, a policy with good aged care facilities would be more attractive to the aged, but the provider must allow the young to choose it if they wish.

On the supply side there are two spatially differentiated, but otherwise identical, service providers/insurers.⁶ There is no price competition, and providers are paid a pre-determined (possibly differentiated) fee per enrollee.⁷ Providers attract (deter) consumers by increasing (decreasing) the quality of the bundles offered, but they must respect incentive compatibility constraints imposed by the requirement that each individual is permitted to select any bundle offered.

We first characterize the socially optimal provision of insurance under both full information and asymmetric information, in which case incentive compatibility constraints must be

⁵ Glazer and McGuire model the insurance industry as perfectly competitive in the sense of Rothschild and Stiglitz (1976). If the purchaser can costlessly observe individuals' types, then the optimal type-based prices just cover the costs of providing the socially optimal health services to each risk type separately. However, if payments to providers can instead be made on the basis of a signal that is only partially informative about an individual's underlying risk type, the optimal risk-adjusted payments either under- or over-pay compared with the optimal type-based prices. Glazer and McGuire's paper is primarily concerned with this statistical issue.

⁶ It would be straightforward to generalize the model to $n > 2$ providers located on a circle as in Salop (1979).

⁷ This payment could be implemented by giving each individual a voucher with a given face value (possibly differentiated across consumers) that can be used only to purchase a bundle of services from one of the providers.

respected. Our simple specification is general enough to admit the possibility of socially optimal insurance policies that differ across individuals. This seems to be an important aspect of the environment that is missing from GM00's seminal contribution, in which the optimal insurance policy is the same for all risk types, but is consistent with the authors' more recent (2002) contribution. In our simple framework, the full information pair of socially optimal insurance policies is never incentive compatible, as high-risk individuals wish to be covered by the policy for low-risks. Under conditions of asymmetric information the policies are thus distorted: the generosity of the policy designed for high-risks is increased – too much is spent on high-risks, and the generosity of the policy for low-risks is reduced – too little is spent on them – and the mix of services offered to low-risks is also distorted – they get the “wrong mix” of services.

Next we derive the equilibrium insurance policies offered by the market, again under conditions of both full and incomplete information. Our assumption about how plans compete is different to that in the recent literature on HMO markets of, e.g., FGM, and GM02. The model of these papers assumes that the level of service quality is defined by setting the marginal benefit of the service equal to a strategically chosen shadow price. This shadow price is the same for all policies offered by the insurer, so that the marginal expected benefit of each service is equated across individuals with different policies. Although this sounds efficient, there are two problems with such an assumption about firm behavior. First, a set of policies characterized by a given vector of shadow prices does not in general maximize the firm's profits, and second the policies offered to different individuals using the shadow pricing rule are not necessarily incentive compatible. In our model, firms choose the components of insurance policies directly (without being constrained to equate marginal expected benefits across individuals) to maximize profits, and when appropriate, these policies are constrained to be incentive compatible.

Having characterized the equilibrium outcome, it is then a relatively simple matter to specify the payments that implement the (full- and information constrained-) social optimum. These optimal type-specific payments cover the costs of providing the socially optimal bundle to each type of individual, plus a mark-up over cost. This mark-up, equal to the unit travel cost parameter, is necessary when providers have some market power and consumers cannot all costlessly switch from one provider to another. When payments must be based on imperfectly informative signals, the type-specific optimal payments need to be adjusted using the linear transformation proposed by GM00.

Biglaiser and Ma (2003) present a model that is close to that of the current paper. As in the current paper, consumers are spatially differentiated, as well as having different risks or types. They consider the performance of different market structures – in particular examining the issue of “carve outs” in health insurance provision – but do not address the question of differential (i.e., risk-adjusted) payments, except tangentially in their conclusion. Similarly, Olivella and Vera-Hernández (2005) model equilibrium in spatially differentiated insurance markets, but focus more on existence issues than risk adjustment. Finally, Jack (2001) considered the nature of equilibrium in a model with spatially differentiated consumers, but under the constraint that insurance providers set a uniform price and allowed individuals to buy as much insurance as they wished at that price. The primary focus of that paper, however, was on the structure of optimal payments based on endogenous signals (in particular, actual spending levels).

In the next section we present a description of consumer preferences and costs in a model with two types of health services and heterogeneous risks. In Section 3 we characterize socially optimal insurance policies under full and incomplete information. Section 4 characterizes market

equilibrium in the presence of spatial consumer heterogeneity with uniform pricing (i.e., no risk adjustment). In Section 5 optimal risk-adjusted payments are derived, first under the assumption that individual characteristics are observable by the purchaser, and then when types are imperfectly observable. Section 6 concludes.

2. Preferences and costs

There are two health care services, x and y , and two possible states of nature, g (good health) and b (bad health). The probability that state b occurs is θ , and the probability that state g occurs is $1 - \theta$. We refer to θ , which can differ across the population, as the “type” of an individual. This is one source of individual heterogeneity; we will introduce a second locational dimension of heterogeneity in Section 4. Service y – thought of as acute care – is only useful in the bad state, and in that state provides benefit $\psi(y)$. Service x – “regular” health care – provides benefits $\xi(x)$ in both states. However, we allow the benefit of service x to be state-dependent, being $\xi_b(x)$ in state b and $\xi_g(x)$ in state g , with $\xi'_b(x) \leq \xi'_g(x)$ —i.e., in the bad state the productivity of regular health care may be lower than in the good state. The expected benefit (compared with no insurance)—of an insurance policy that provides a bundle of services $P = (x, y)$ is

$$u(x, y; \theta) = \zeta(x, \theta) + \theta\psi(y). \tag{1}$$

If the benefit of service x is state-independent, then $\zeta(x, \theta) = \xi(x)$. This is the case considered by GM00. Otherwise, $\zeta(x, \theta) = \theta\xi_b(x) + (1 - \theta)\xi_g(x)$.

To permit the derivation of simple expressions for optimal and equilibrium outcomes, from time to time we shall place further restrictions on the benefit functions, and assume that $\psi(y) = \ln y$, $\xi_g(x) = \alpha \ln x$, and $\xi_b(x) = \beta \ln x$, with $\alpha > \beta$. The expected utility of a bundle $P = (x, y)$ then simplifies to

$$u(x, y; \theta) = \gamma(\theta) \ln x + \theta \ln y$$

where $\gamma(\theta) = \alpha(1 - \theta) + \beta\theta$.

The expected expenditure associated with an insurance plan (x, y) is denoted

$$m(x, y; \theta) = x + \theta y.$$

We define the maximal benefit of a policy with expected expenditure m as

$$v(m; \theta) = \max_{x,y} u(x, y; \theta) \quad \text{s.t.} \quad x + \theta y = m.$$

For a given θ , the bundle $(x(m), y(m))$ that solves this problem defines a functional relationship between x and y . Write this as $x = \hat{x}(y; \theta)$. In the case of log utility, $\hat{x}(y; \theta) = \gamma(\theta)y$.

Any plan $(\hat{x}(y; \theta), y)$ provided to a θ -type individual maximizes her expected utility given the expected cost of the policy, and is efficient in this sense. For a given value of θ any efficient policy is determined solely by the value of y chosen. Fig. 1 illustrates the loci of efficient plans in the log utility case for two types of individuals, $\theta_H > \theta_L$. (Note that $\gamma'(\theta) = -(\alpha - \beta) < 0$. Of course, when $\alpha = \beta$, the two lines coincide.)

With log utility, the expected cost of an efficient policy with acute care level y is

$$\begin{aligned} c(y; \theta) &= \gamma(\theta)y + \theta y \\ &= \hat{\gamma}(\theta)y \end{aligned} \tag{2}$$

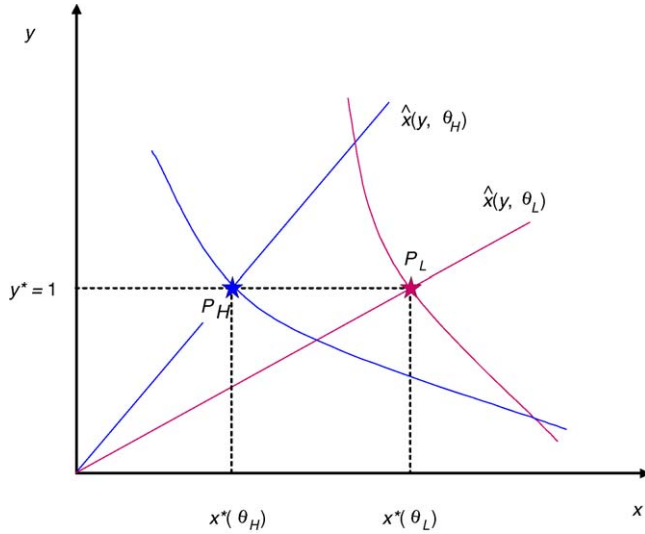


Fig. 1. The line $\hat{x}(y; \theta_H)$ is the locus of plans that provide an individual of type θ_H with maximal expected utility for a given level of expected cost. Similarly, $\hat{x}(y; \theta_L)$ is the locus of such plans for θ_L types, where $\theta_L < \theta_H$. The curved lines are indifference curves for each type. At the surplus maximizing plans, P_L and P_H , both types have $y^* = 1$, but in general $x^*(\theta_H) < x^*(\theta_L)$.

where $\hat{\gamma}(\theta) \equiv \theta + \gamma(\theta)$. If a plan’s benefit levels are increased through a marginal increase in the quality of the acute service, matched by an efficient increase in the quality of regular care, the incremental cost is $c_1(y; \theta) = \hat{\gamma}(\theta)$, constant for given θ . We assume that $\alpha - \beta < 1$, so that the marginal cost of expanding plan quality is higher for individuals with greater risk: $\hat{\gamma}'(\theta) = 1 - (\alpha - \beta) > 0$. (Note that this is trivially satisfied in the case that $\alpha = \beta$.)

Finally, with log utility the expected utility earned by a θ -type individual insured with an efficient plan with acute service quality y is

$$\begin{aligned} \omega(y, \theta) &= u(\hat{x}(y; \theta), y, \theta) \\ &= \Gamma(\theta) + \hat{\gamma}(\theta) \ln y \end{aligned} \tag{3}$$

where $\Gamma(\theta) = \gamma(\theta) \ln \gamma(\theta)$. The marginal utility of increased plan quality is $\omega_1(y, \theta) = \hat{\gamma}(\theta)/y$, which is increasing in θ under the assumption that $\alpha - \beta < 1$, i.e., $\omega_{12} = (1 - (\alpha - \beta))/y > 0$. That is, higher risk individuals value increased plan quality more than lower risk individuals.

3. Socially optimal plans

In this section of the paper and the next we assume that θ can take on one of two values, $\theta_L < \theta_H$, and that the proportion of θ_L -types is $\rho \in (0, 1)$. We first characterize the pair of socially optimal plans when plans can be assigned to individuals as a function of their types. We then characterize incentive compatible socially optimal plans. We note that the introduction of locational heterogeneity across consumers in Section 4 will not alter the socially optimal plans (under either full or incomplete information), since we will assume then that travel costs are linear and additively separable from the benefit, $u(\cdot)$, derived from insurance.

3.1. Full information

The expected surplus associated with a plan (x, y) covering a θ -type individual is

$$V(x, y; \theta) = u(x, y; \theta) - m(x, y; \theta).$$

The surplus maximizing plan must be efficient (in the sense above), and is therefore characterized by $y^*(\theta)$ that solves

$$\omega_1(y; \theta) = c_1(y; \theta). \tag{4}$$

With log utility, this reduces to

$$y^* = 1,$$

and the surplus maximizing level of regular care is

$$\begin{aligned} x^*(\theta) &= \hat{x}(1, \theta) \\ &= \gamma(\theta). \end{aligned}$$

When the value of regular care is state-independent, $x^* = \alpha$, independent of θ . The case of state-dependent benefits is shown in Fig. 1. The level of expected utility of each type at the optimum is denoted

$$\omega_L^* = u(x_L^*, y_L^*; \theta_L) \quad \text{and} \quad \omega_H^* = u(x_H^*, y_H^*; \theta_H). \tag{5}$$

3.2. Incentive compatibility

When $\alpha = \beta$, the optimal plans for both high- and low-risk individuals are identical, and hence trivially incentive compatible. However, it is clear from Fig. 1 that when $\alpha > \beta$, all individuals – both high- and low-risk – will prefer the plan P_L to plan P_H . This is simply because both plans promise the same quality of acute care, but P_H provides lower quality regular care than P_L .⁸

To characterize the incentive compatible socially optimal pair of plans (P'_L, P'_H) we anticipate the “no-distortion at the top” result that P'_H will be efficient, that is

$$P'_H = (\hat{x}(y_H, \theta_H), y_H)$$

for some y_H . Also, from Fig. 1 it is clear that any incentive compatible pair of plans must provide lower quality acute care to low-risk individuals, $y_L < y_H$. Thus, for any value y_H , and for any $y_L < y_H$ define $\tilde{x}(y_L, y_H)$ as satisfying

$$\omega(y_H, \theta_H) = u(\tilde{x}, y_L, \theta_H).$$

The bundle (\tilde{x}, y_L) gives the same expected utility to a θ_H -type individual as the efficient policy with acute care service quality y_H . $\tilde{x}(y_L, y_H)$ is the highest quality of regular care that can be provided under a plan for low-risk individuals in conjunction with acute care quality y_L that will ensure that high-risk individuals prefer the efficient plan designed for them, (\hat{x}, y_H) .

⁸ This result is not as asymmetric as it might seem. High risk individuals use acute care more often than low risks, while both types use regular care equally often. Thus, the expected quality of acute care is higher for high risks than for low risks.

The expected cost of the plan (\tilde{x}, y_L) is

$$\tilde{c}(y_L, y_H; \theta_L) = \tilde{x}(y_L, y_H) + \theta_L y_L \tag{6}$$

which satisfies the properties⁹

$$\frac{\partial \tilde{c}}{\partial y_L} < 0 \quad \text{and} \quad \frac{\partial \tilde{c}}{\partial y_H} > 0.$$

Similarly, the expected utility earned by a low-risk individual insured by the plan (\tilde{x}, y_L) is

$$\tilde{\omega}(y_L, y_H; \theta_L) = u(\tilde{x}, y_L; \theta_L) \tag{7}$$

which satisfies¹⁰

$$\frac{\partial \tilde{\omega}}{\partial y_L} < 0 \quad \text{and} \quad \frac{\partial \tilde{\omega}}{\partial y_H} > 0.$$

To find the optimal pair of incentive compatible plans, we solve

$$\max_{y_L, y_H} \rho[\tilde{\omega}(y_L, y_H; \theta_L) - \tilde{c}(y_L, y_H; \theta_L)] + (1 - \rho)[\omega(y_H, \theta_H) - c(y_H, \theta_H)].$$

If (y'_L, y'_H) solves this problem, then the optimal plans are $P'_L = (\tilde{x}(y'_L, y'_H), y'_L)$ and $P_H = (\hat{x}(y'_H), y'_H)$. The first order conditions for (y'_L, y'_H) are

$$\theta_L - \text{types} : \tilde{\omega}_1(y_L, y_H; \theta_L) = \tilde{c}_1(y_L, y_H; \theta_L) \tag{8}$$

and

$$\theta_H - \text{types} : \omega_1(y_H, \theta_H) = c_1(y_H, \theta_H) - \left(\frac{\rho}{1 - \rho} \right) [\tilde{\omega}_2(y_L, y_H; \theta_L) - \tilde{c}_2(y_L, y_H; \theta_L)]. \tag{9}$$

The square-bracketed term in (9) is positive, so the incentive compatible social optimum is characterized by excessive but efficient spending for high-risk individuals, and sub-optimal and inefficient spending on low-risks. That is,

$$y'_H > y_H^* \quad \text{and} \quad x'_H = \hat{x}(y'_H, \theta_H)$$

$$y'_L < y_L^* \quad \text{and} \quad \hat{x}(y'_L, \theta_L) < x'_L < x_L^*$$

This result is illustrated in Fig. 2 for the case of log utility. The full-information socially optimal plans P_L and P_H are shown as before. Also shown is an iso-surplus curve for low-risks. This

⁹ With log utility, the expressions for these derivatives are

$$\frac{\partial \tilde{c}}{\partial y_L} = \theta_L - \theta_H \left(\frac{y_H}{y_L} \right)^{(1 + (\theta_H/\gamma_H))} \quad \text{and} \quad \frac{\partial \tilde{c}}{\partial y_H} = \gamma_H \left(\frac{y_H}{y_L} \right)^{(\theta_H/\gamma_H)}$$

which can be seen to have the reported signs. The marginal effects have these signs as long as the indifference curves in Fig. 1 satisfy the single crossing property.

¹⁰ With log utility, the expressions for these derivatives are

$$\frac{\partial \tilde{\omega}}{\partial y_L} = \theta_L \left(1 - \frac{\gamma_L}{\gamma_H} \frac{\theta_H}{\theta_L} \right) < 0 \quad \text{and} \quad \frac{\partial \tilde{\omega}}{\partial y_H} = \gamma_L \left(1 + \frac{\theta_H}{\gamma_H} \right) > 0$$

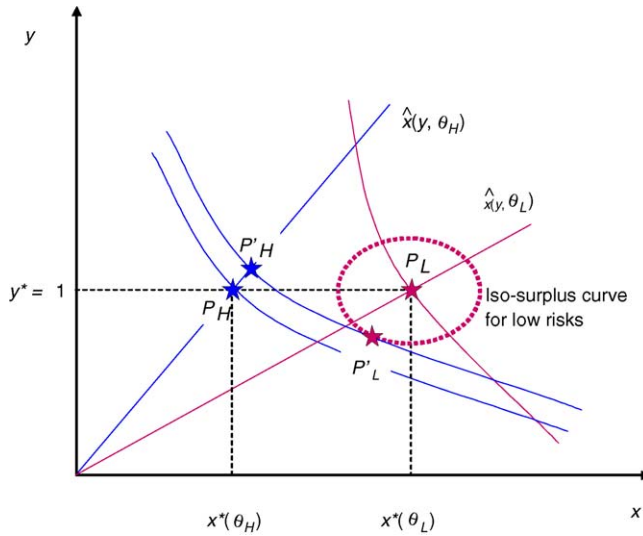


Fig. 2. Socially optimal incentive compatible plans, P'_L and P'_H .

curve is flat whenever $x = x^*(\theta_L)$, it is vertical whenever $y = y^* = 1$, and along the efficiency locus $\hat{x}(y, \theta_L)$ it has the same slope as a θ_L -type indifference curve (which is steeper than the θ_H -type indifference curve through the same point). The optimal pair of incentive compatible plans, P'_L and P'_H must be on the same θ_H -type indifference curve, and P'_L must maximize low-risk surplus subject to being on such an indifference curve. Thus, in order to induce high-risks to choose the plan designed for them, both plans are distorted. First, the generosity of the plan for high-risks is increased, although it remains efficient (P'_H is to the north-east of P_H , but on the $\hat{x}(y, \theta_H)$ locus); second, the generosity of the plan for low-risks is reduced, and the mix of services provided under this plan is distorted away from acute care (P'_L is to the south-west of P_L , and to the right of the $\hat{x}(y, \theta_L)$ locus).

Finally, we define ω'_L and ω'_H to be the utility earned by θ_L - and θ_H -types, respectively, at the social optimum:

$$\omega'_L = u(x'_L, y'_L; \theta_L) \quad \text{and} \quad \omega'_H = u(x'_H, y'_H; \theta_H). \tag{10}$$

From Fig. 2 it is clear that high-risks earn higher expected utility at the social optimum under asymmetric information, and low-risks earn lower expected utility. Also, expected expenditure is higher for high-risks, and lower for low-risks. Note that the lower expenditure on low-risks is *in spite of the fact* that the policy they are allocated, P'_L , does not provide the given (lower) level of expected utility at minimum cost (i.e., P'_L is not on the curve $\hat{x}(y, \theta_L)$). Even though the policy is distorted towards service x , the quality of this regular care is still lower than at the full information optimum, $x^*(\theta_L)$.

4. Market provision

The government’s objective is to implement socially optimal insurance coverage by paying private providers on behalf of individuals who enroll in their plans. Equivalently, the government endows each individual with a voucher that can be used to purchase insurance from a private

provider. In this paper we assume a horizontally differentiated market structure, and introduce a second dimension of consumer heterogeneity. A number of papers on risk adjustment have adopted such an approach, including Biglaiser and Ma (2003), GM02, Jack (2001), and Olivella and Vera-Hernández (2005). GM00, in which consumers differ only by risk type, is a notable exception. Incorporating a horizontal or spatial dimension of heterogeneity into the model allows us to investigate providers' incentives to expand or contract the size of the market they serve in a meaningful sense. This in turn means that risk-adjusted payments not only can serve the purpose of altering the allocative properties of a Rothschild–Stiglitz equilibrium in the presence of adverse selection as in GM00, but also can be used to reduce providers' incentives to deter participation of certain individuals.¹¹

Individuals of each risk type are distributed uniformly on the interval $[0, 1]$. There are two providers with identical cost functions, one at $\delta=0$ and the other at $\delta=1$. A θ_J -type individual ($J=L, H$) at position $\delta \in [0, 1]$ who consumes a bundle (x, y) delivered by provider 0 has net utility

$$U^0(x, y, \delta; \theta_J) = u(x, y; \theta_J) - t\delta$$

where t is a unit travel cost. If she acquires such a bundle from provider 1 her net utility is

$$U^1(x, y, \delta; \theta_J) = u(x, y; \theta_J) - t(1 - \delta).$$

As long as $u(x, y; \theta_J) > t\delta$, a J -type individual at δ will prefer to be served by provider 0 than to go without the services. We assume that t is small enough that the optimal utilities defined by (5) or (10) are sufficiently high as to make it socially optimal for all individuals to be served. That is, universal coverage of the population is welfare maximizing.

To start with, we assume the providers are paid a single price or fee, r , per individual to whom they provide services (alternatively, the value of the voucher is the same for all individuals), which is financed through non-distortionary taxation, the distribution of which is not modelled explicitly here.¹² As in GM00 and GM02, individuals pay nothing directly to the provider from whom they obtain insurance. Provider i announces a pair of policies, (P_i^L, P_i^H) , where $P_i^J = (x_i^J, y_i^J)$, and serves all individuals who request either policy.¹³

Our assumption about the way in which providers choose insurance policies is different to that made in some recent models of HMO behavior such as FGM, and GM02. In those models, providers are assumed to choose a vector of shadow prices q which is used to ration health care across individuals. Adapting the GM02 model to our environment, the marginal expected benefit of service x provided to each covered individual is equated to the shadow price q_x chosen by

¹¹ In a perfectly competitive model, expected profits per participant are zero in equilibrium, so providers are indifferent to expansions or contractions in demand on the extensive margin. This indifference seems at odds with the practical concerns of health insurance providers.

A related advantage of the model over that of perfect competition, helpfully suggested by a referee, is that the equilibrium in our model is more robust. In a Rothschild–Stiglitz equilibrium *all* individuals are indifferent between purchasing insurance from all providers. If an insurer can actively discriminate between risk types at some small cost, then it can induce all the bad risks to switch to another insurer. However, in a model with horizontal heterogeneity, a small amount of active discrimination will induce only a small number of switches, since most consumers are *not* indifferent between providers.

¹² If travel costs are assumed to be pecuniary, then our model specification effectively assumes the marginal utility of income is unity for all individuals, so the distribution of the tax burden has no impact on social welfare.

¹³ We assume that providers cannot make insurance policies contingent on an individual's location.

the insurer, thereby equalizing the marginal expected benefit of this service across the plan's clients.¹⁴

However, as noted by Ma (2003), such a strategy is in general not profit maximizing: an insurer cares not about Pareto efficiency, but about profits. In addition, we show below that such a strategy does not necessarily yield an incentive compatible pair of policies, so cannot constitute an equilibrium. To begin, we analyze the equilibrium under complete information.

4.1. Full information equilibrium

We assume providers have full information about risk types, and can perfectly discriminate. The two providers effectively compete in two markets separately—the θ_L -market and the θ_H -market. In the θ_j -type market, if provider 0 offers a policy that provides gross utility ω_j^0 and provider 1 offers one providing gross utility ω_j^1 , then as long as $\omega_j^0 + \omega_j^1 \geq t$ all θ_j -type individuals will participate by choosing a bundle from one or other of the providers. The proportion who choose firm 0's bundle is $1/2[1 + (\omega_j^0 - \omega_j^1)/t]$, and the proportion served by firm 1 is $1/2[1 + (\omega_j^1 - \omega_j^0)/t]$.

Because the firms face no incentive compatibility constraints they have no reason to offer a policy that is not efficient. That is, the policies chosen by firm 1 (say) take the form $P_j^1 = (\hat{x}(y_j^1, \theta_j), y_j^1)$ for some y_j^1 . The insurer can attract more θ_j -type consumers by increasing the quality of services offered to them, but faces no incentive to distort the mix of services. Thus, firm 1's problem is to choose values of y_L^1 and y_H^1 to maximize its profits, taking as fixed firm 0's choices of y_L^0 and y_H^0 . Then firm 1's aggregate profit from offering (y_L^1, y_H^1) is

$$\begin{aligned} \pi_1(y_L^1, y_H^1; y_L^0, y_H^0) = & \rho[r - c(y_L^1, \theta_L)] \left(\frac{1}{2} + \frac{\omega(y_L^1, \theta_L) - \omega(y_L^0, \theta_L)}{2t} \right) \\ & + (1 - \rho)[r - c(y_H^1, \theta_H)] \left(\frac{1}{2} + \frac{\omega(y_H^1, \theta_H) - \omega(y_H^0, \theta_H)}{2t} \right). \end{aligned}$$

Firm 0's profit is defined similarly. Given firm 0's policy choice in the θ_j -market, firm 1's best response satisfies the first order condition

$$[r - c(y_j^1, \theta_j)]\omega_1(y_j^1, \theta_j) = c_1(y_j^1, \theta_j)[t + \omega(y_j^1, \theta_j) - \omega_j^0]. \quad (11)$$

At a symmetric equilibrium, $y_j^1 = y_j^0 \equiv y_j^E$, characterized by

$$\omega_1(y_j, \theta_j) = \frac{tc_1(y_j, \theta_j)}{[r - c(y_j, \theta_j)]}. \quad (12)$$

In general, and in contrast to models that assume firms compete by choosing shadow prices, the marginal expected benefit of each service is not necessarily equated across individuals of different

¹⁴ In the first part of GM02 there is no uncertainty, and individuals have heterogeneous preferences. Shadow prices are then used to allocate quantities of different health services across the individuals. In the second part of the paper (as detailed in the appendix thereto), there is uncertainty, and shadow prices are used to allocate services *ex post*. Thus, an insurance policy, under uncertainty, promises different levels of each service in different health states. The model of the current paper, although it is in an environment of uncertainty, is formally closer to the first part of GM02, since an insurance policy provides given levels of the two kinds of medical care that are *the same* across health states—i.e., there is no *ex post* rationing of care. It is for this reason that we think of the components of the insurance policy— x and y —as measuring the quality of physicians and/or facilities the plan employs. The insured individual enjoys the same quality of care independent of the health state.

risk types. To see this most easily, consider the example of log utility, in which case condition (12) becomes

$$\omega_1(y_J, \theta_J) = \frac{t\hat{\gamma}_J}{r - \hat{\gamma}_J y_J^E},$$

and the equilibrium quality of acute care provided by the market to θ_J -type individuals is

$$y_J^E = \frac{r}{t + \hat{\gamma}_J} \tag{13}$$

where $\hat{\gamma}_J = \hat{\gamma}(\theta_J)$. The corresponding quality of regular care is $\hat{x}(y_J^E, \theta_J)$. Considering, for example, service y , in equilibrium the ratio of the marginal expected benefit of acute care provided to θ_L -types to that provided to θ_H -types is

$$\frac{u_2(x_L^E, y_L^E, \theta_L)}{u_2(x_H^E, y_H^E, \theta_H)} = \frac{\theta_L}{\theta_H} \left(\frac{t + \hat{\gamma}_L}{t + \hat{\gamma}_H} \right) < 1.$$

since $\hat{\gamma}_H > \hat{\gamma}_L$.¹⁵

Finally, note from (13) that the quality of acute care is lower for high-risks than for low-risks. These results hold even when the benefit of regular medical care is state-independent, i.e., even when $\alpha = \beta$. However, recall our observation above that any incentive compatible pair of policies must provide lower quality acute care to low-risk individuals. It is immediately clear that under asymmetric information condition (12) cannot characterize the equilibrium policies provided by the market.

4.2. Equilibrium under incentive constraints

If insurers cannot discriminate between risk types, the policies they offer must be incentive compatible. To characterize the equilibrium outcome we employ the cost and utility functions defined in (6) and (7). When firm 0 offers a pair of incentive compatible policies that yield gross utilities ω_L^0 and ω_H^0 for low- and high-risks, respectively, firm 1’s aggregate profit from offering the pair of policies $\{(\tilde{x}(y_L^1, y_H^1), y_L^1), (\hat{x}(y_H^1, \theta_H), y_H^1)\}$ is

$$\begin{aligned} \pi_1(y_L^1, y_H^1; y_L^0, y_H^0) = & \rho[r - \tilde{c}(y_L^1, y_H^1; \theta_L)] \left(\frac{1}{2} + \frac{\tilde{\omega}(y_L^1, y_H^1; \theta_L) - \omega_L^0}{2t} \right) \\ & + (1 - \rho)[r - c(y_H^1, \theta_H)] \left(\frac{1}{2} + \frac{\omega(y_H^1, \theta_H) - \omega_H^0}{2t} \right). \end{aligned}$$

We denote the quality of regular and acute care at the symmetric Nash equilibrium by x^{EJ} and y^{EJ} . These values are characterized by the following conditions for the quality of acute care provided to low- and high-risks, respectively:

$$\theta_L\text{-types} : \tilde{\omega}_1 = \frac{t\tilde{c}_1}{[r - \tilde{c}]} \tag{14}$$

¹⁵ Footnote 14 explains why we compare marginal *expected* benefits, and not marginal benefits in each state of nature.

and

$$\theta_H - \text{types} : \omega_1(y_H^{E'}, \theta_H) = \frac{tc_1(y_H^{E'}, \theta_H)}{[r - c(y_H^{E'}, \theta_H)]} - \frac{\rho}{(1 - \rho)} \left(\frac{(r - \tilde{c})\tilde{\omega}_2 - t\tilde{c}_2}{r - c(y_H^{E'}, \theta_H)} \right) \tag{15}$$

where \tilde{c} and its derivatives are evaluated at $(y_L^{E'}, y_H^{E'}; \theta_L)$.

By incorporating the incentive constraint into the insurers' maximization problems, we have guaranteed that the policies offered in equilibrium will be incentive compatible, and in particular that they will satisfy the monotonicity condition $y_L^{E'} < y_H^{E'}$. We note that if firms compete by equating marginal expected benefits to strategically set shadow prices (as suggested in footnote 14), then the purported equilibrium is not necessarily incentive compatible. To see this, consider the case of log utility again. Equating marginal expected benefits of each service across the two risk types means that the insurance policies (x_L, y_L) and (x_H, y_H) satisfy

$$\frac{x_L}{x_H} = \frac{\gamma_L}{\gamma_H} \quad \text{and} \quad \frac{y_L}{y_H} = \frac{\theta_L}{\theta_H}. \tag{16}$$

Such a pair of policies is incentive compatible if and only if

$$u(x_L, y_L; \theta_L) \geq u(x_H, y_H; \theta_H)$$

and

$$u(x_H, y_H; \theta_H) \geq u(x_L, y_L; \theta_L).$$

Using log utility, substituting (16), and adding the two inequalities implies that the condition

$$(\gamma_L - \gamma_H) \ln \left(\frac{\gamma_L}{\gamma_H} \right) - (\theta_H - \theta_L) \ln \left(\frac{\theta_H}{\theta_L} \right) \geq 0$$

must hold if the policies are to be incentive compatible. This simplifies to

$$(1 - \beta) \geq \frac{\ln(\theta_H/\theta_L)}{\ln(\gamma_L/\gamma_H)}. \tag{17}$$

The left hand side of this condition is less than one, so if $\theta_H/\theta_L > \gamma_L/\gamma_H$ (both of which are greater than one) then the pair of policies cannot satisfy (17), i.e., they cannot be incentive compatible. It is straightforward to show that in fact, as long as $\alpha > (1 - \beta)(\theta_L + \theta_H)$, this is the case. Note that this is a sufficient, but not necessary, condition for insurance policies set using shadow prices to fail to be incentive compatible.

5. Risk adjustments

In the model of market provision above, it was assumed that providers of insurance were paid a fee r per person served by the government. Equivalently, all individuals received a voucher with the same face value that could be used to purchase insurance. The question of risk adjustment concerns the equilibrium and welfare impacts of paying different fees for people with different characteristics—that is, differentiating the face value of a voucher on the basis of such characteristics. For example, if the government could pay different fees for θ_H - and θ_L -types it might be able to induce the market to provide the welfare maximizing bundles. We examine this possibility in Section 5.1. Even if such full information is not available to the government, it might be able to profitably differentiate between individuals on the basis of a partially informative signal

(GM00). That is, even if the government cannot identify the underlying risk a given individual faces, characteristics such as age, gender, race, etc. may be correlated with that underlying risk. The statistical adjustment to the type-based risk-adjusted prices is considered in Section 5.2.¹⁶

We first note that even if providers have full information regarding individuals' types, the government may wish to differentiate payments by type, simply because the equilibrium in each type-specific market may not necessarily correspond to the social optimum. That is, risk adjustment is not necessarily a response to adverse selection. Rather, the costs of providing socially optimal bundles to different consumer types may well differ, and if prices are not correspondingly differentiated, the bundles delivered in equilibrium will be sub- (or super-) optimal. When providers have incomplete information about individuals' types, incentive compatibility constraints distort the optimal bundles, changing their costs and the prices necessary to induce their delivery.

5.1. Risk adjustments based on underlying risk

In this sub-section we assume that the government has full information about risk types, and can differentiate the fee paid to providers on the basis of an individual's type. (In the following subsection we assume the government has a partially informative signal about risk types.) However, since the government relies on the market to provide insurance, it must implicitly choose amongst market equilibria. If the market also has full information about risk types, then the full social optimum can be implemented by a suitable design of the fee structure. However, if the market operates under conditions of asymmetric information, or indeed if risk-based discrimination by insurers is prohibited, then the government can only hope to implement an equilibrium outcome that respects incentive compatibility constraints.¹⁷

One approach to optimal risk adjustment is to search for a pair of payments, (r_H, r_L) , that maximizes welfare under the constraints imposed by the information structure faced by the market. However, as long as the revenues that finance these payments are raised in a lump-sum fashion, the welfare maximizing pair of payments yields the same outcome as that which would be achieved by the central planner directly choosing optimal bundles *under the same information constraints*. That is, the optimal pair of type-specific fees, (r_H, r_L) , implements the utility levels, and corresponding insurance policies, associated with either (4), or (8) and (9), depending on whether the market has full information or not. We consider each case in turn.

5.1.1. Optimal risk adjusters when the market has full information

For given payments r_H and r_L , the full information Nash equilibrium satisfies the trivial generalization of (12),

$$\omega_1(y_J, \theta_J) = \frac{tc_1(y_J, \theta_J)}{[r_J - c(y_J, \theta_J)]} \quad (18)$$

¹⁶ In this paper, and much of the literature, the signals upon which risk adjusters are based are exogenous. Jack (2001) examines how payments should be related to endogenous medical expenditures under conditions of moral hazard.

¹⁷ Of course, this is a mildly contrived exercise: if the government had such complete information to make type-specific payments to providers, it could simply inform the providers so they no longer faced any IC constraints. More realistic is a model in which the government receives an observable, exogenous, signal that is correlated with an individual's type, and can base the payment to a provider on this. See Glazer and McGuire (2000) and Section 5.2 below.

for $J=L, H$. The full information social optimum defined by (4) requires $\omega_1(y_J, \theta_J) = c_1(y_J, \theta_J)$, so the fee for θ_J -types should be set to satisfy

$$r_J^* = c(y_J^*, \theta_J) + t.$$

That is, under full information, the risk-adjusted payment for each type should consist of the cost of efficiently providing the optimal bundle, plus the per unit travel cost. In the case of log utility, this is simply

$$r_J^* = \hat{y}(\theta_J) + t.$$

Why is the risk-adjusted price higher than the cost of the optimal insurance? The answer is simply that if the price was set at $r_J = c(y_J^*, \theta_J)$, then a firm that offered the optimal policy would earn zero profits on all its θ_J -type clients. Reducing the quality of insurance below the optimum would increase profits on all inframarginal consumers, at the loss of market share corresponding to an elastic response by marginal consumers, thereby increasing total profits above zero. The higher is the travel cost parameter t , the lower is this elasticity of demand, and the more profitable quality reductions will be. Thus, the risk-adjusted price needs to exceed the cost of the optimal insurance policy by a certain margin (t), and the less competitive the market, the higher this margin needs to be.

5.1.2. *Optimal risk adjusters when the market has incomplete information*

When the market has incomplete information, or is prohibited from explicitly using the information it has, about individuals' risks, the government can only hope to implement the incentive compatible social optimum, with quality of acute care y'_L and y'_H given by (8) and (9). If the risk-based payments provided by the government are r_L and r_H , the market equilibrium satisfies the generalizations of (14) and (15):

$$\theta_L\text{-types} : \tilde{\omega}_1 = \frac{t\tilde{c}_1}{[r_L - \tilde{c}]} \tag{19}$$

and

$$\theta_H\text{-types} : \omega_1(y_H^{E'}, \theta_H) = \frac{tc_1(y_H^{E'}, \theta_H)}{[r_H - c(y_H^{E'}, \theta_H)]} - \frac{\rho}{(1 - \rho)} \left(\frac{(r_L - \tilde{c})\tilde{\omega}_2 - t\tilde{c}_2}{r_H - c(y_H^{E'}, \theta_H)} \right). \tag{20}$$

The optimal payment for low-risks (recalling that at the social optimum $\tilde{\omega}_1 = \tilde{c}_1$) is now

$$r'_L = \tilde{c}(y'_L, y'_H, \theta_H) + t,$$

and that for high-risks satisfies

$$\begin{aligned} r'_H &= c(y'_H, \theta_H) + t \left[1 - \left(\frac{\rho}{1 - \rho} \right) \right] \left(\frac{\tilde{\omega}_2(y'_L, y'_H, \theta_L) - \tilde{c}_2(y'_L, y'_H, \theta_L)}{\omega_1(y'_H, \theta_H)} \right) \\ &= c(y'_H, \theta_H) + t \end{aligned}$$

substituting for $(\tilde{\omega}_2 - \tilde{c}_2)$ from (9). Once again, the optimal payment for each type covers the cost of providing the socially optimal policy for that type, plus the unit travel cost parameter to account for the strength of competition.

5.2. Risk adjustments based on signals of risk

GM00 make the observation that if payments are made on the basis of an imperfectly informative signal of an individual’s type, then it is optimal to over-pay for some signals and underpay for others. First, following GM00, suppose there is just a single binary signal, $s \in \{0, 1\}$, and that q_J is the probability that a θ_J -type individual has signal 1. The signal is informative if $1 > q_J > q_H \geq 0$, so that θ_H -types have a higher chance of receiving the signal $s = 1$. If a payment R_s is made on behalf of individuals with signal s , then the average amounts earned on θ_H - and θ_L -types are

$$\bar{r}_H = q_H R_1 + (1 - q_H) R_0 \tag{21}$$

and

$$\bar{r}_L = q_L R_1 + (1 - q_L) R_0, \tag{22}$$

respectively. Assuming providers cannot select based on the signal itself,¹⁸ \bar{r}_J is the expected payment made to a firm serving a θ_J -type individual. Denote the optimal type-based payment by \hat{r}_J , where

$$\hat{r}_J = \begin{cases} r_J^* & \text{if the market has full information} \\ r'_J & \text{if the market has incomplete information.} \end{cases}$$

If providers are risk neutral and maximize expected profits, then in order to implement the optimal bundles, we choose the R_s so that the expected payment for a θ_J -type coincides with the optimal type-based payment, i.e., $\bar{r}_J = \hat{r}_J$ for $J=L, H$. Setting the right hand sides of (21) and (22) equal to \hat{r}_H and \hat{r}_L , respectively, and solving for \hat{R}_0 and \hat{R}_1 , we find

$$\hat{R}_0 = \frac{q_H \hat{r}_L - q_L \hat{r}_H}{q_H - q_L}$$

and

$$\hat{R}_1 = \frac{(1 - q_L) \hat{r}_H - (1 - q_H) \hat{r}_L}{(q_H - q_L)}.$$

These are the same as GM00’s Eqs. (16) and (17). Thus, the method of statistically correcting the optimal payments is unchanged, but the “conventional” payments (the r^J) are adjusted, to both account for the strength of competition between firms (as measured by the travel cost parameter) and the composition of the optimal insurance policies.

5.2.1. Multiple signals

In practice, there are likely to be multiple signals (age, gender, family history, etc.) that are correlated with an individual’s type. In addition, these signals need not have binary support. In general, suppose there are n signals, and signal i has a support Σ_i . Let $s = (s_1, \dots, s_n)$ be a

¹⁸ This assumption is made by Glazer and McGuire (2000), who allude to open enrollment requirements in US health plans and non-discrimination rules. However, its validity is questioned by Newhouse (2002, page 197). Recalling the discussion in paragraph 6 of the introduction, we believe our assumption is consistent with both positions. Newhouse notes that if providers are paid little for enrolling children, then they will have an incentive to provide low quality pediatric services. This is precisely the assumption of the current paper—providers choose policies given the reimbursement rates set by the purchaser. However, they are *not* permitted to deny coverage to (families with) children because the reimbursement rates they attract are low (or for any other reason).

vector of signals, where $s_i \in \Sigma_i$, and suppose $f_J(s)$ is the probability density function of θ_J -types on $\Sigma = \Sigma_1 \times \Sigma_2 \times \dots \times \Sigma_n$. Also, let $R(s)$ be the payment received by a provider if it delivers services to an individual with a vector of signals s . The average payment made to a θ_J -type individual is

$$\bar{r}_J = \int_{\Sigma} R(s) f_J(s) ds.$$

If \hat{r}_J is the optimal type-based payment for θ_J -types, then any function $R(s)$ for which $\bar{r}_J = \hat{r}_J$ for each J implements the social optimum.

For example, consider the case of $n = 1$ but when Σ_1 is a closed interval in \mathbb{R} . There are two natural specifications for $R(s)$: first, if

$$R(s) = \begin{cases} 0 & \text{if } s < s_0 \\ r_0 & \text{if } s \geq s_0 \end{cases} \tag{23}$$

then the payment to a provider for an individual is zero if her signal is sufficiently low and a positive constant if it is above a threshold. Alternatively, let

$$R(s) = \beta_0 + \beta_1 s \tag{24}$$

so the payment is affine in the signal. Since each specification has two degrees of freedom, except in pathological cases it will be possible to find a payment schedule within either class that satisfies $\hat{r}_J = \bar{r}_J$ for $J=L, H$, where r_J is the optimal type-based payment.¹⁹ For example, using (24), $\bar{r}_J = \beta_0 + \beta_1 \mu_J$, where μ_J is the expected value of the signal for θ_J -types.

For $n > 2$, even within a given class of function, there is no longer a unique payment schedule. For example, the affine version becomes

$$R(s) = \beta_0 + \sum_{i=1}^n \beta_i s_i \tag{25}$$

which has $n + 1$ degrees of freedom.²⁰ If for each risk type the signals are independent, i.e., $f_J(s) = f_{J1}(s_1) \times \dots \times f_{Jn}(s_n)$, then the average payment on behalf of θ_J -types is

$$\bar{r}^J = \beta_0 + \sum_{i=1}^n \beta_i \mu_{Ji}. \tag{26}$$

Any values of β_0, \dots, β_n that solve $\bar{r}_J = \hat{r}_J$ for $J=L, H$ implement the social optimum. GM02 resolve the indeterminacy when the number of signals is larger than the number of types by

¹⁹ One might be induced to choose the smoother version (24) than the step function (23) if the assumption regarding providers' inability to select based on the signal is not considered robust.

²⁰ An alternative specification of $R(s)$ is

$$R(s) = \begin{cases} 0 & \text{if } z \cdot s < s_0 \\ r_0 & \text{if } z \cdot s \geq s_0 \end{cases}$$

where z is a vector of weights that sum to unity, and s_0 is a scalar-valued threshold parameter. Under this payment scheme, the payment to a provider is zero for individuals with signals that are, on (weighted) average, sufficiently low, and equal to s_0 for those with a (weighted) average signal above the threshold.

choosing a solution vector that also has a minimum variance property.²¹ Alternatively, one might assume providers are not risk neutral, or that at some cost they can select on the basis of the signal, and from these fundamentals derive optimal payment schedules.²²

6. Conclusions

We have used a model that incorporates both spatial heterogeneity and risk heterogeneity to examine the features of optimal prices paid by a purchasing agency on behalf of individuals with different underlying characteristics. One contribution of the paper is to specify which insurance policies are in fact optimal, and how much they cost. Indeed, consumer heterogeneity implies that a single insurance policy (e.g., uniform national insurance) is not necessarily optimal. We have shown that, assuming providers cannot explicitly differentiate between individuals (i.e., select directly on the basis of risk type), then decentralizing the delivery of the (constrained) optimal insurance policies is possible as long as providers are paid risk-adjusted fees for each individual they serve. When the payment can be made on the basis of an individual's underlying risk type, it should be sufficient to cover the cost of delivering the socially optimal policy for that person, plus a mark-up over cost. This mark-up, equal to the unit travel cost parameter, accounts for the fact that by lowering quality below the socially optimal level, providers can increase profits on inframarginal consumers, for whom switching to an alternative provider is costly. While the optimal policy, and its cost, varies by type, the mark-up is independent of the type. Thus, optimal payments are differentiated by an amount that is independent of the strength of competition.

If payments can be made only on the basis of a partially informative signal, the optimal risk-based payments should be adjusted according to a simple linear transformation, identified by GM00. If there are more signals than types, some indeterminacy arises regarding the coefficients of this transformation. An important empirical question to be addressed is the estimation of the probability densities linking types and signals.

Acknowledgements

Thanks to Ian Gale and Marius Schwartz for helpful discussions, and to Joe Newhouse, Tom McGuire, and two referees for thoughtful comments.

References

- Biglaiser, G., Ma, Ching-to Albert, 2003. Price and quality competition under adverse selection: market organization and efficiency. *RAND Journal of Economics* 34 (2), 266–286.
- Eggleston, K., 2000. Risk selection and optimal health insurance-provider payment systems. *Journal of Risk and Insurance* 67, 175–198.
- Frank, R., Glazer, J., McGuire, T., 2000. Measuring adverse selection in managed health care. *Journal of Health Economics* 19, 829–854.
- Glazer, J., McGuire, T., 2000. Optimal risk adjustment of health insurance premiums: an application to managed care. *American Economic Review* 90 (4), 1055–1071.
- Glazer, J., McGuire, T., 2002. Setting health plan premiums to ensure efficient quality in health care: minimum variance optimal risk adjustment. *Journal of Public Economics* 84 (2), 153–175.

²¹ In fact, the condition in their paper is that the number of signals is larger than the number of goods offered in each bundle. In our model, this is equal to the number of types of consumers.

²² Eggleston (2000) assumes incentives to select expensive patients are effectively limited by medical ethics.

- Jack, W., 2001. Controlling risk selection incentives when health insurance contracts are endogenous. *Journal of Public Economics* 80, 25–48.
- Ma, Ching-to Albert, 2003. Managed care and shadow price. *Health Economics Letters* 7 (3), 17–20.
- Murgolo, Maggie, 2002. MCBS highlights: comparison of Medicare Risk HMO and FFS enrollees. *Health Care Financing Review* 24 (1), 177–185.
- Newhouse, J., 1996. Reimbursing health plans and health providers: efficiency in production versus selection. *Journal of Economic Literature* 34, 1236–1263.
- Newhouse, J., 2002. *Pricing the Priceless: A Health Care Conundrum*. MIT Press, Cambridge MA.
- Olivella, P., Vera-Hernández, M., 2005. Competition among differentiated health plans under adverse selection, unpublished manuscript.
- Rothschild, M., Stiglitz, J., 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics* 90, 629–649.
- Salop, S., 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10, 141–156.
- Shen, Y., Ellis, R.P., 2002. Cost minimizing risk adjustment. *Journal of Health Economics* 21 (3), 515–530.
- Van de Ven, W., Ellis, R., 2000. Risk adjustment in competitive health plan markets. In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*. North-Holland.