Altruism and Incentives in Public and Private Health Care[1]

Ching-to Albert Ma[2]

October 2004

[1]An earlier version was presented at the Yrjö Jahnsson Foundation 50^{th} Anniversay Symposium on Incentives and Finance of the Health Care System, August 9-10, 2004. My thanks to Hans Keiding and Klaus Kultti for their comments.

[2]Department of Economics, Boston University, 270 Bay State Road, Boston, Massachusetts 02215, USA; <ma@bu.edu>. Many ideas in this paper have been developed jointly with Gary Biglaiser of the University of North Carolina, and Philippe Choné of CREST-LEI, Paris.

Abstract

I hypothesize that physicians have diverse preferences. Some are dedicated and provide services at high quality despite the lack of incentive or monitoring; some others value patients' benefits. I then set up two models to consider the interaction between the public and private healthcare sectors. In the first model, physicians may be allowed to refer patients in the public system to their private practice. In the second, the public sector offers low-powered incentive schemes to attract physicians. In each case, I demonstrate that more altruistic and dedicated physicians work in the public sector, while physicians who value profits more work in the private market.

1 Introduction

In many countries health care is by provided by firms in the public and private sectors. The effect of interaction between the public and private sectors on physician incentives is the subject of this chapter. Providers of health care in public and private sectors are often subject to different incentives. While the market mechanism is expected to work in the private market, the public system is typically characterized by simple and low-powered incentives. For the public sector, a conventional model, relying on straightforward utility or profit-maximizing preferences for economic agents, is inappropriate. Such a model would predict uniformly poor service quality and work effort there. This prediction is certainly inconsistent with casual observations; despite the lack of incentives, tasks and services are still being performed in the public sector.

The key hypothesis of this paper is that some healthcare providers are either sincere or have altruistic preferences. This deviation from the conventional wisdom may be less readily appreciated by academic economists than practitioners. Academic economists tend to believe that simple, consistent maximizing behavior is enough to study most social institutions. I believe that it is a misguided hypothesis for the health sector. It is indeed true that many models predict that maximizing behaviors may actually lead to sincere and altruistic actions. They typically push the "rational" preferences one step higher into the hierarchy of analysis; agents behave sincerely because of their fear for punishment or their expectation of future rewards. I tend to believe that this method obscures rather than enlightens. For policy matters, it is more important to study the practical implications of sincere and altruistic behaviors rather than simply to push the question one step back.

In this paper, I consider models in which healthcare providers may be sincere and altruistic. I assume that physicians consist of heterogenous groups. Some are more sincere and altruistic than others. I study how a public system can exploit this heterogeneity. Who among the heterogenous physicians will work in the public system, who in the private market?

Healthcare reforms seek to enhance efficiency, reduce cost, and maintain access and quality. Healthcare providers play the crucial role in this discussion. How they behave and react against changes in the public and private systems drive many of the conclusions in an analysis of any reform. In this paper, some physicians behave like the conventional "economic man," maximizing their utilities or profits. Others either do not exploit their private information or supply services

with good qualities even when monitoring and incentives are absent.

My basic conclusion is that public policy should steer sincere and altruistic physicians to work in the public system, while the marketplace disciplines those who seek personal gains. It is a straightforward conclusion due to the use of low-powered incentives in the public system. Allowing the self-interest seeking providers the opportunity of higher profits in the private sector actually alleviates the inefficiency in the public sector.

I present two models. In the first, the quality of health services in the public system is unmonitored. There are sincere physicians who will supply services at good qualities even when there is no explicit incentive. The profit-maximizing physicians will supply services at a minimal quality. The public system may introduce a referral system, allowing physicians to transfer patients in the public system to their own private practice. The profit-maximizing physicians will exploit this market opportunity when they and their patients can mutually gain. The referral system may save costs in the public sector because the government does not pay for the costs of care of patients referred to the private market.

In the second model, physicians are altruistic, their preferences being combinations of their profits and patient benefits. There is a continuum of altruistic characteristics. Only the physicians know their preferences; the government or firms in the private sector do not. I first present the optimal mechanism that firms in the private sector will offer to these altruistic physicians, who value the services provided to their patients. The optimal mechanim respects the physicians' private information, distorting quantities as well as giving information rent to physicians.

I then let the public sector offer low-powered incentives to attract some of the physicians into public service. In effect, the public system selects the more altruistic physicians, who value the services provided to their patients.. This has also the effect of changing the profile of characteristics of physicians remaining in the private sector. The overall effect may tend to reduce information cost in the private sector.

Some papers have studied dual job incentives in the health sector, see Rickman and McGuire (1999), and Gonzalez (2004). The paper by Gonzalez studies incentives of excessive treatment in the public sector when physicians there may signal their abilities to the private market. The distortion for excessive treatment because a physician performing well in the public system may generate a higher private demand. Other papers have looked at that same issue generally, see Che

(1995) and Lewis and Sappington (1989) for example. Che considers the incentives to invest in human capital when a regulator can seek post-tenure appointment at a regulated firm. Lewis and Sappington studied incentive designs when the agent's interests are in many environments. Papers on the multi-task principal-agent problem deal with similar issues, see Holmstrom and Milgrom (1994) for a general model and Ma (1994) for a study on the health sector. The general approach assesses the effectiveness of incentive mechanisms when agents have a lot of instruments available, but they adopt the usual assumption that agents maximize their own personal gains.

Many papers have examined the effect of the private sector on waiting lists in the public sector. Iversen (1997) considers a dynamic model of rationing by waiting lists; see also Barros and Olivella (2002). Besley, Hall, and Preston (1998) show that waiting lists in the UK National Health Services are positively associated with private insurance. Ma (2003) considers cost incentives when the public system rations supplies.

Recent advances in economic theory have incorporated ideas from the psychology literature. My hypothesis of heterogenous behaviors regarding profit and altruistic acts among physicians builds on this literature. Rabin (1993) studies a formal game-theoretic model on fairness. In Alger and Ma (2003) some physicians reveal information truthfully despite the lack of incentives to do so. Jack (2004) uses a model of altruistic providers to study cost and quality incentives. Frank (1996) argues that some workers are willing to give up high wages to serve in companies that provide social services. Finally, Besley and Ghathak (2003) let firms adopt social missions to attract workers who prefer to work in these firms. Brekke, Kverndokk and Nyborg (2003) propose a model of moral motivation, which is supported by agents' preferences for following an endogenous social norm.

The next section presents the two models. The following two then contain the analysis. Concluding remarks are in the last section.

2 Models of Dual Job Incentives and Physician Agency

I present in this section two models of physician behaviors. I call the first one a dual job model; the second, a physician agency model. Common to both models is a set of consumers; they are indexed by a parameter α, which is positive and follows a distribution F with density f on the positive support $[\underline{\alpha}, \overline{\alpha}]$. The parameter α measures a patient's valuation of medical services. The benefit that the consumer obtains when service q is provided by a physician is $\alpha B(q)$, where B is

a strictly increasing and concave function. The distribution of α captures diverse preferences as well as illness severities. It is assumed that the value of α is only known to the patient and the physician, not to the insurer, regulator or government.

The cost of service q, borne by a physician, is given by the strictly increasing and convex function $C(q)$. The physician may receive reimbursement to cover this cost. The measure of medical service can be interpreted either as quantity or quality. I adopt an interpretation of quality for q in the dual job model, and an interpretation of quantity in the physician agency model. I use different assumptions on whether the cost of medical service is verifiable; details follow soon.

I write down a benchmark for efficiency. For a given value of α, the service that maximizes the net benefit is given by

$$q(\alpha) \equiv \arg\max_q \alpha B(q) - C(q), \qquad \text{or} \qquad \alpha B'(q(\alpha)) = C'(q(\alpha)).$$

I call this the first-best level of service. The first best refers to the situation where the value of α is common knowledge. The first-best service $q(\alpha)$ is increasing in α. The rest of the paper is concerned with asymmetric information, when α is only known to the patient and the physician.

There is a set of physiciains, and I normalize the total number of physicians so that each doctor treats one and only one patient. There are always more patients than physicians, so that no physician has to be idle. In the dual job model, there are only two types of physicians. The first type I call the dedicated doctors; these are physicians who follow recommendations and work in a sincere fashion. I will let the regulator or government issue an instruction on the provision of medical service in the public sector. So the dedicated doctors will simply follow this instruction; of course, the regulator will have to reimburse them for the costs. The second type of physicians are common economic agents who maximize their utility or profits. Let the fraction of dedicated doctors among all physicians be μ.

In the second model, there is a continuum of physician types. Each physician's preferences depend on a combination of his own profit and his patient's benefit. The precise weight in the combination of physician profit and patient benefit varies across the population of physicians. So some physicians have preferences that put more weights on patient benefits than others. I let the parameter β represent this weight; this parameter follows a distribution G with density g on the positive support $[\underline{\beta}, \overline{\beta}]$. If a physician of type β is paid R after having supplied services q to a patient, then his preferences are given by $R - C(q) + \beta B(q)$. The physician gets to observe the

patient's preference parameter α. Both α and β are the physician's private information.

There are two sectors: the public and the private. All consumers are insured and may get services from the public sector. I assume that consumers pay a constant user fee for services in the public sector, and this is normalized to zero.[1] A private sector for medical services may also exist. Also, if the government permits it, a physician working in the public sector may participate in the private sector too. In the public sector, high-powered incentive mechanisms are not used. As I demonstrate later, this does not imply that the public sector becomes nonviable. My hypothesis is that the public sector uses low-powered incentives to attract some types of providers.

The market mechanism in the private sector is assumed to facilitate the exchange between consumers and physicians, or to make available incentive mechanisms to solve problems due to asymmetric information. For the dual job model, in the private sector, consumers and physicians can freely contract on quality and payment. Most important is the physician-consumer coalition's ability to tailor quality q to the specific consumer preference parameter α. For the physician agency model, I assume that in the private sector an incentive mechanism can be used to extract the physician's private information of α and β.

In each of the two models, I hypothesize unconventional behaviors for healthcare providers. In the dual job model, the dedicated doctors behave in a nonstrategic way. In the physician agency model, doctors have altruistic preferences (although a doctor's degree of altruism is his private information). Both assumptions deviate from the usual and conventional one of pure profit or utility maximization. However, they are important for a study of public services. Strong incentive systems in the public sector are seldom found. Yet, in most countries projects and tasks in the public sector are being carried out. A conventional assumption of complete self-interest would imply that nothing can be achieved in the public sector. This is incompatible with casual observations.

I now describe the various specific assumptions on information, variable verification, and the extensive form for each model.

[1]The public system integrates insurance and provision. For insurance purposes, consumers do not bear the full marginal cost of service.

2.1 Dual Job Incentives

For the dual job model, I first consider a benchmark in which all physicians work in the public sector. Then the government has to decide if the physicians are allowed to work in the private sector. I assume that quality q is nonverifiable. So the government can only pay a fixed renumeration to a physician who works in the public system. Treating a patient in the public system results in the physician receiving a fixed wage, R. The government will also issue an instruction on the quality that should be supplied to patients in the public sector. Let this be q^r. The instruction is supposed to inform the physicians the level of quality that should be supplied, and the level of compensation R is set at $C(q^r)$ which covers the associated cost of the recommended quality.

The dedicated doctors and profit-maximizing physicians react differently against the fixed wage and quality recommendation. The dedicated doctors follow the instruction, supply services at quality q^r, and accept the compensatory wages. The profit-maximizing physician supply services at a minimum quality, $\underline{q}$, because no explicit incentives are present in the public sector; these agents then net a profit of $C(q^r) - C(\underline{q})$.

If participation in the private market is allowed, a profit-maximizing physician can refer a patient from the public sector to the private sector. After a referral, the government will not have to pay the physician. In the private sector, the physician will bargain and contract with the patient. I assume that the bargaining is efficient. This, for example, is a property of the Nash bargaining solution. Efficiency implies that the patient will obtain quality $q(\alpha)$. I assume that the physician and the patient will then split the surplus by an appropriate transfer.

For simplicity, I assume that the dedicated doctors will not participate in the private market. The purpose of this model is to demonstrate how opening up the private market will alleviate the inefficiency in the public sector, so I will not analyze explicitly change of behavior by the dedicated doctors.

2.2 Physician Altruism

For the physician agency model, I assume that q measures healthcare quantity, and is verifiable, and the payment R can be contingent on how much the physician has chosen to supply. I first consider a benchmark in which all physicians work in the private sector. The government then considers hiring some of these physicians to work in the public system.

The physician agency model is one of multi-dimensional uncertainty. The physician gets to observe the patient's valuation parameter α and his own degree of altruism β. I assume that a firm in the private sector will use an optimal mechanism. This takes the form of a direct revelation mechanism with truthful revelation being an equilibrium. A mechanism takes the form of a pair of functions: $(R(\alpha,\beta), q(\alpha,\beta))$. The physician (who knows the true values of α and β) is asked to make a report. If he reports values of α' and β' he is paid $R(\alpha',\beta')$ and required to supply quantity $q(\alpha',\beta')$. Truthful revelation as an equilibrium means that the functions R and q guarantee that misreporting values of α and β is suboptimal.

I require that the physician makes nonnegative profits. Although the physician's utility depends on both profits and consumer benefit, it is unlikely that he can sustain a monetary loss. This assumption is much more natural than a reservation utility constraint.

I will present the optimal mechanism in the private sector. When the government attempts to hire physicians, it maintains its low-powered incentive system, simply offering a wage in exchange for the physician providing some quantities. The purpose of this model is to study whether more or less altruistic phyciaisns find working in the public system more attractive.

2.3 Policies

I will regard permitting public physicians to work in the private sector, and expanding the public sector as policy choices. The question remains what guides the policy decisions. I use a utilitarian approach: the government is supposed to adopt a policy that maximizes the sum of consumer benefits less any payments made by consumers and the government. In practice a government's agenda may not be so entirely reflected by a utilitarian approach. Nevertheless, it is a useful welfare index to start with. For a more realistic discussion, I will mention distribution issues in the analysis.

3 Dedication and Dual Job Incentives

In this section, I study the dual job model. I begin with the benchmark when physicians in the public sector are unable to take up jobs in the private sector. A patient is randomly matched to a physician in the public sector. Let the government recommend a quality q^r and offers a payment of $R = C(q^r)$ to a physician for providing a treatment to a patient. The dedicated doctors will supply services at quality q^r; they earn zero profits. The profit-maximizing doctors will merely

provide a service at minimum quality $\underline{q}$; they receive payment $C(q^r)$, hence each netting a profit $C(q^r) - C(\underline{q})$. The welfare index consists of the expected consumer (net) benefit less the payment made by the government

$$\mu \mathrm{E}(\alpha) B(q^r) + (1-\mu)\mathrm{E}(\alpha) B(\underline{q}) - C(q^r), \tag{1}$$

where $\mathrm{E}(\alpha)$ is the expected value of α. In my setup, the government uses a low-powered incentive system, paying all physicians a fixed wage. It relies on the goodwill of the dedicated doctors to maintain the recommended quality to some consumers. The payments to the profit-maximizing doctors are just transfers. From (1), one can characterize the optimal recommended quality by the first-order condition with respect to q^r: $\mu \mathrm{E}(\alpha) B'(q^r) = C'(q^r)$. Because the payments to the profit-maximizing doctors do not yield any quality, the marginal benefit is only proportional to the fraction of dedicated doctors (μ).

Next, I consider the policy of allowing the physicians to participate in the private market. Referring a patient to a private practice is now possible. I assume that the dedicated doctor does not participate in the private market; only the profit-maximizing doctors do. I use a Nash bargaining solution to describe the outcome in the private sector. First, the disagreement point is one where a physician and the patient fail to agree on a referral. In this case, the patient's utility is $\alpha B(\underline{q})$, while the physician's utility is $C(q^r) - C(\underline{q})$. An agreement is a pair (p, q), where p is the price the consumer pays the physician and q the quality the physician has to provide. A Nash bargaining solution maximizes the joint surplus (above the disagreement) and specifies a price to split the surplus in some ratio, say s, $0 < s < 1$, between the two parties.

The quality in the Nash bargaining solution maximizes the joint surplus $\alpha B(q) - C(q)$; and the quality level is the first-best $q(\alpha)$. The physician and patient will be able to agree on a referral when their joint surplus is higher than in the disagreement point. This is the condition: $\alpha B(q(\alpha)) - C(q(\alpha)) \geq \alpha B(\underline{q}) - C(\underline{q}) + C(q^r)$, which is satisfied whenever α is above a certain threshold, say $\widehat{\alpha}$. If α is below this threshold, the physician and patient cannot agree on a referral with mutual benefits; they stay within the public system.

In summary, if the consumer is matched with a dedicated physician, she receives treatment at quality q^r in the public sector. Otherwise, a referral is recommended by the physician if her valuation parameter α is sufficiently high. In that case, she receives the efficient quality $q(\alpha)$, and pays the physician a price, say $p(\alpha)$. If a referral is not recommended, the patient stays within the

public system, receives the minimum quality $\underline{q}$ but pays nothing.

I now write the welfare index in this regime:

$$\mu\left\{\mathrm{E}(\alpha)B(q^r)-C(q^r)\right\}+(1-\mu)\left\{\int_{\underline{\alpha}}^{\widehat{\alpha}}[\alpha B(\underline{q})-C(q^r)]f(\alpha)\mathrm{d}\alpha+\int_{\widehat{\alpha}}^{\overline{\alpha}}[\alpha B(q(\alpha))-p(\alpha)]f(\alpha)\mathrm{d}\alpha\right\}.$$

The first term refers to the welfare generated by the dedicated physicians. The second term is divided into two parts: the first refers to the net benefit provided by the profit-maximizing doctors in the public sector, while the second is the surplus from a referral to the private sector. To compare the welfare index under the referral regime with that in (1), I write the price in the private market as a markup over cost: $p(\alpha)\equiv C(q(\alpha))+m(\alpha)$. So now I rewrite the welfare index as:

$$\begin{gathered}\mu\left\{\mathrm{E}(\alpha)B(q^r)-C(q^r)\right\}+\\(1-\mu)\left\{\textstyle\int_{\underline{\alpha}}^{\widehat{\alpha}}[\alpha B(\underline{q})-C(q^r)]f(\alpha)\mathrm{d}\alpha+\int_{\widehat{\alpha}}^{\overline{\alpha}}[\alpha B(q(\alpha))-C(q(\alpha))-m(\alpha)]f(\alpha)\mathrm{d}\alpha\right\},\end{gathered}$$

which is further simplifed to

$$\begin{gathered}\mu\mathrm{E}(\alpha)B(q^r)+(1-\mu)\mathrm{E}(\alpha)B(\underline{q})-C(q^r)+\\(1-\mu)\textstyle\int_{\widehat{\alpha}}^{\overline{\alpha}}\left\{[\alpha B(q(\alpha))-C(q(\alpha))-m(\alpha)]-[\alpha B(\underline{q})-C(q^r)]\right\}f(\alpha)\mathrm{d}\alpha.\end{gathered}\tag{2}$$

The welfare index in (2) embeds a number of properties. By the definition of Nash bargaining, $\alpha B(q(\alpha))-C(q(\alpha))-m(\alpha)>\alpha B(\underline{q})$, for all $\alpha>\widehat{\alpha}$, so for the same recommended quality q^r, the welfare index is higher than (1) when the physicians are allowed to refer patients to the private market. There are two sources of the gain. The first is due to the efficiency of quality (as a function of the valuation parameter α) since the physician and the patient can contract on quality and price in the private market. Second, there is some cost saving: the last term $(1-\mu)(1-F(\widehat{\alpha}))C(q^r)$ is the government's cost saving from not paying those physicians who refer patients.

The private market allows an opportunity for the profit-maximizing doctors to supply higher quality to those consumers with higher valuations. It also alleviates the incentive problem in the public sector. A profit-maximizing doctor forgos the rent $C(q^r)-C(\underline{q})$ in order to refer a patient to the private practice. The private sector serves as an efficient supply and a sorting mechanism.

Although the welfare index must be improved when referral is allowed, it does not imply that *all* consumers must be better off. This depends on whether the government actually changes the recommended quality level once referral is permitted. There are two counteracting effects for the government to consider. First, private market referral leads to some cost savings. The

government's payment $C(q^r)$ is more likely given to a dedicated doctor. Hence there is a tendency for the government to raise the recommended quality.

Second, private market referral leads to qualities that are efficient relative to consumers' valuations. The recommended quality in the public sector applies to all consumers who seek services there; consumers with different valuations receive the same quality. Some of the rigidity can be avoided when referrals are possible; consumers in the private sector receives qualities that are more in-line with their valuations. To exploit this market flexibility, the government may want to steer more consumers to the private sector. The way to achieve this is to lower the recommended quality level in the public sector; more consumers will then find the private sector attractive.

If the recommended quality actually increases because the cost savings are so significant, then all consumers become better off compared to the regime where referrals are impossible. Clearly those who stay behind in the public sector are better off. Those who decide to accept the referral must also find it beneficial too. If the recommended quality falls because the government wants to steer more consumers to the private market, then those consumers who stay behind in the public sector become worse off.[2]

3.1 Robustness and Adverse Effects

The gain in welfare when the profit-maximizing physicians are allowed to refer patients is based on the trading opportunity in the private sector. This is a fundamental point. I have used the Nash bargaining solution, which implies an efficient allocation from a referral. This allows me to make the point in a succinct fashion, but the precise efficiency property in the private market is unimportant. What is important is that the opportunity for trading in the private sector does not require the patient or the physician to forgo any surplus in the public sector. For the Nash bargaining solution, this is captured by the disagreement point being defined as the allocation the patient and the physician would have achieved if referral were impossible.

The welfare property can be supported even when there is asymmetric information between the patient and the physician, or when some other market frictions exist. These obstacles will reduce the potential gain, as well as the volume of trade in the private market. However, in any equilibrium, if the physician succeeds in making a referral, both the physician and the patient must

[2]The efficiency gain in the private sector is strong enough to outweigh the reduction of consumer benefit in the public sector.

expect to gain from the decision. For referral to be a rational outcome, each party must expect to be better off than staying within the public sector. Market frictions affect how much each party gains, and how often referrals occur. Nevertheless, as long as there is some likelihood of a successful referral, the expected welfare must increase.

Adverse and unintended effects must be considered when the private market is opened.[3] I have assumed that when referral to the private market is allowed, this is the only new activitiy for profit-maximizing doctors. In practice, new activities or effects may occur. First, this may affect the behaviors of the dedicated doctors. The government enjoys the sincere behaviors of the dedicated doctors in the public system. If these doctors find the higher income from the private market attractive, they may abandon the public service. Morale may be affected, and incentive costs may increase.

Private market participation may further adversely affect profit-maximizing doctors' behavior in the public sector. They may decide to lower quality at the public sector even more. This let them devoate more time and energy to private patients. Moreover, a lower quality at the public sector may make patients more willing to accept a referral suggestion. This deterioration of quality in the public sector typically affects the less well-off consumers.

There may be some countervailing incentives, however. A physician may have an incentive to perform well in the public system in order to signal to the private market his effort or ability (Gonzalez, 2004). The asymmetric information may come from physicians having different abilities; their performance may reveal some information. Signaling however may lead to other distortions. For example, in order to enhance his performance in the public sector, patients may be treated excessively. There is no reason, however, to expect any signaling incentive to exactly counteract incentives for other adverse reactions.

The supply of physicians in the public sector may be affected when the private market expands. Some physicians may choose to leave the public sector altogether. The departure of the profit-maximizing physicians poses less of a problem for the public system. In any case, they are supposed to take up the profit opportunties there. The departure of the dedicated physicians is a definite negative for the public sector.

Adverse effects due to the expansion of the private market may be mitigated by regulations.

[3]See Biglaiser and Ma (2003) for more detailed discussions.

If the magnitude of the adverse effect is positively related to the profit opportunities in the referral market, then price ceiling regulation may be considered. This regulation simply says that a physician's referral fee cannot exceed a certain amount. In practice, this may be implemented by a limit on physician income. For example, a regulation may forbid a physician working in the public system to have more than 30% of his income derived from private work. The lower the ceiling, the less a physician can exploit the profit opportunities in the private sector. There is then less incentive to engage in activities that will further reduce quality in the public system.

4 Altruism and Physician Agency

In this section, I study the physician agency model. I begin with the optimal mechanism for physicians in the private sector. Then I study the impact of the public sector offering low-powered incentives to hire some of these physicians. The public and private sectors are in effect sorting physicians with different degrees of altruism. Recall that the altruistic physician's utility is given by $R - C(q) + \beta B(q)$. Firms are supposed to be competitive; they offer incentive mechanisms to physicians in order maximize consumers' expected utility, subject to incentive and nonnegative profit constraints for the physicians.

Recall that the mechanism $(R(\alpha, \beta), q(\alpha, \beta))$ must satisfy the truth-telling (or incentive compatibility) and nonnegative profit constraints. The incentive constraints are:

$$(\alpha, \beta) = \arg\max_{\alpha', \beta'} \quad R(\alpha', \beta') - C(q(\alpha', \beta')) + \beta B(q(\alpha', \beta')).$$

The right-hand side expression is the utility a physician who has altruism parameter β and who is matched with a consumer with valuation parameter α. The incentive constraint says that the physician can do no better than reporting truthfully this information. The maximum or indirect utility is $\max_{\alpha', \beta'} \; R(\alpha', \beta') - C(q(\alpha', \beta')) + \beta B(q(\alpha', \beta'))$.

Because the physician's preferences are not directly dependent on the patient's valuation parameter α, the indirect utility function is only a function of β. Let the indirect utility be $U(\beta)$. Again, because the physician's preferences are independent of α, any incentive mechanism is generically independent of α, the patient's valuation.[4] In other words, the physician's private information

[4] The requirement for the genericity is due to the following. The indirect utility function U is the maximum of affine linear functions of β. So it must be convex in β, and hence almost everywhere differentiable. If $B(q(\alpha, \beta))$ were not independent of α, the differentiability of U would be violated.

about α cannot be extracted directly. So from now on I write a mechanism as $(R(\beta), q(\beta))$.

I can now write the indirect utility function as $U(\beta) = R(\beta) - C(q(\beta)) + \beta B(q(\beta))$. By the Envelope Theorem, $U'(\beta) = B(q(\beta))$. Furthermore, because U is the maximum of affine linear functions of β, it is convex. So $U''(\beta) = B'(q(\beta))q'(\beta) \geq 0$; this implies that any quantity q in an incentive compatible mechanism must be nondecreasing in β. In fact incentive compatibility is equivalent to U being convex, and q being nondecreasing.

Next, I consider the nonnegative profit condition. Let $\pi(\beta) \equiv R(\beta) - C(q(\beta))$. I use the definition of the indirect utility function to obtain $\pi(\beta) \equiv U(\beta) - \beta B(q(\beta) = U(\beta) - \beta U'(\beta)$. Differentiating the profit function, I get

$$\frac{\mathrm{d}\pi}{\mathrm{d}\beta} = \frac{\mathrm{d}}{\mathrm{d}\beta}[U(\beta) - \beta U'(\beta)] = -\beta U''(\beta) = -\beta B'(q(\beta))q'(\beta) \leq 0. \tag{3}$$

The profit is nonincreasing with respect to the physician's altruism parameter.

The properties of the quantity and profit as functions of the physician's altruism private information are atypical. These two functions actually have opposite monotonicity characteristics. In fact, a firm hiring a physician with private information on altruism (as well as consumer valuation) must give up information rent. This rent is obtained by the physician through two components, the payment R and the quantity q. Nevertheless, a more altruistic physician is willing to give up more profits in return for a higher quantity (and hence higher consumer benefit). So a more altruistic physician earns less profit.

The profit decreases as the degree of altruism increases, at a rate determined by the quantity increase rate; see (3). Both the quantity and profit schedules have to be compatible, the former being increasing while the latter decreasing, with (3) relating the rates they can change. If the quantity increases too rapidly, then the profit falls to zero and the nonnegative profit constraint binds. This then forces the quantity to become constant. Actually, if the profit constraint binds at $\widehat{\beta}$, then it binds for all $\beta > \widehat{\beta}$, and as a result the quantity becomes constant too. Altruism and nonnegative profits may lead to a choice of a set of quantities that are insensitive to the variation of the altruism parameter when β is sufficiently high.

4.1 The Optimal Mechanism for Altruistic Physicians

I assume that firms in the private sector operate in a competitive environment. So they choose mechanisms to maximize consumers' expected utility. The objective function is

$$W \equiv \int \int [\alpha B(q(\beta)) - R(\beta)]\ h(\alpha, \beta)\ \mathrm{d}\alpha\ \mathrm{d}\beta,$$

where h is the joint density of α and β. Because the physician's private information about α cannot be extracted, I rewrite this objective fucntion as follows:

$$W \equiv \int [\alpha_m(\beta) B(q(\beta)) - R(\beta)]\ g(\beta)\ \mathrm{d}\beta,$$

where $\alpha_m(\beta)$ is the conditional mean of α given β. The optimal mechanism is one that maximizes this objective function subject to the incentive and nonnegative profit constraints.

I will skip over the derivation of the optimal mechanism; see Chone and Ma (2004) for the technical details. It is, however, important to mention a key step. The optimal mechanism must consider a pooling regime. For a range of higher values of β the quantity becomes constant (no longer strictly increasing in β), and profit zero. For lower values of β, the nonnegative profit constraint does not bind, and quantity is strictly monotone increasing in β. So the objective function for the optimal mechanism can be written in two parts. For low values of β, it is a separating regime; physicians with different degrees of altruism receive different payments and provide different quantities. For high values of β, it is a pooling regime; physicians with different degrees of altriusm receive the same payment and provide the same quantity. Moreover, physicians in the pooling regime earn zero profit.

I will use the following assumption: $\alpha_m(\beta) + \beta + G(\beta)/g(\beta)$ is continuous and nondecreasing in β. This assumption guarantees that the optimal quantity is indeed nondecreasing. Furthermore, I assume that the unconditional mean of α is less than $\alpha_m(\underline{\beta}) + \underline{\beta}$, which guarantees that the pooling regime does not extend to the entire support of β so that some separation remains.

The optimal quantity is depicted in Figure 1. The quantity is continuous, strictly increasing for β between $\underline{\beta}$ and $\widehat{\beta}$, but constant between $\widehat{\beta}$ and $\overline{\beta}$. Let the pooling quantity be denoted by $q(\widehat{\beta})$. The physician profit is illustrated in Figure 2. The profit is continuous, strictly decreasing for β between $\underline{\beta}$ and $\widehat{\beta}$, but remains at 0 between $\widehat{\beta}$ and $\overline{\beta}$. For pooling regime, the indirect utility $U(\beta)$ is $\beta B(q(\widehat{\beta}))$ since profit is zero.

Let me provide some intuition behind the properties of the optimal mechanism. The information about α cannot be extracted directly; it can only be inferred from the information of β revealed by the physician. The conditional expectation $\alpha_m(\beta)$ then becomes a critical part in the determination of the optimal quantity. The hazard rate $G(\beta)/g(\beta)$ takes into account the information rent available to the physician due to his superior information. Furthermore, when the physician's non-negative profit constraint does not bind, part of the information rent is derived from the quantity, which is distorted upward.

For small values of β, profits contribute more to the indirect utility. As quantity begins to increase along with β, the altruistic part $\beta B(q)$ contributes more to the indirect utility. Pooling must occur for high values of β. If it was separating for all values of β, then the quantity would be strictly increasing. Now reducing the quantity for the highest value of β would lead to a second-order loss in terms of quantity efficiency but a first-order gain in terms of information-rent saving. This implies that pooling must be part of the optimal schedule. The assumption that α is less than $\alpha_m(\underline{\beta}) + \underline{\beta}$ guarantees that pooling at the lowest value of β is not as attractive as extracting the information. In this case, the quantity schedule is the one in Figure 1.

Finally, the indirect utility is $U \equiv \pi(\beta) + \beta B(q(\beta))$. Let me assume here that the benefit function B is positive.[5] Figure 3 presents the (convex) indirect utility, the physician's payoff when both profit and altruistic components are taken into account. Here, although physicians who have lower values of β enjoy more profits, their maximum utility is actually lower, since there is less satisfaction from servicing their patients with the lower quantities. Furthermore, for higher values of β, the indirect utility is linear in β since profit is zero: $U = \beta B(q(\widehat{\beta}))$.

4.2 Low-powered Incentives in the Public Sector

I have presented the optimal mechanism in the private sector. Now a government intends to expand its public health services, and must hire some physicians from the private market. I maintain the assumption that the incentive system in the public sector is low-powered. The payment mechanism is simply a fixed wage, in exchange for the physician providing some quantity. Let the quantity chosen by the government be $\widetilde{q}$. The compensation for providing this quantity is the physician's

[5] The benefit function is an ordinal measure; its sign has no relevance for the design of the optimal mechanism. For the derivation of the optimal mechanism, only the marginal benefit B' is relevant. Nevertheless, the sign of the benefit function does affect the sign of the indirect utility, although, in a symmetric fashion, only the sign of U' (which is positive) has a bearing on the optimal mechanism.

cost $C(\widetilde{q})$. Under this scheme, a physician with altruism parameter β obtains a utility $\beta B(\widetilde{q})$.

When the public sector offers this package, how do physicians respond? A physician with altruism parameter β compares $U(\beta) = \pi(\beta) + \beta B(q(\beta))$ in the private sector with $\beta B(\widetilde{q})$ in the public sectors. In Figure 4, I have drawn these two utilities as functions of β. The solid line is the same as in Figure 3 (where the indirect utility from the optimal mechanism is depicted), while the dotted line is $\beta B(\widetilde{q})$. When $\widetilde{q}$ is greater than $q((\widehat{\beta}))$, the utility function for the public sector intersects that for the private sector from below. Those physicians with a degree of altruism higher than $\widetilde{\beta}$ will find it more attractive to work in the public sector.

The simple policy of asking physicians to provide in the public sector a fixed (and higher) quantity attracts those who are more altruistic. A policy of the public sector prescribing a *higher* level of health care than in the private sector is unusual; in most countries, the opposite is true. This is due to the restriction that physicians are paid only their costs of services. A low-powered incentive may also involve a salary, or a lump-sum payment, say T. In this case the utility from joining the public service becomes $T+ \beta B(\widetilde{q})$. Various combinations of T and $\widetilde{q}$ can be considered. The salary component has the effect of raising the level of the dotted line in Figure 4. It is then possible for the government's quantity prescription to be reduced below the highest in the private sector. Figure 5 illustrates an example of this policy. There, the requested quantity in the public sector is lower than the highest in the private sector; in fact, the salary is not sufficiently high to attract those phsyicians who value consumer benefits the most.

The point is that quantity standards and salary schemes in the public sector attract different types of physicians. Depending on the specific kind of health services and providers, the government's offer of salary and standard of care may vary. Each combination sets up a different partition of physicians with different degrees of altruism across the public and private sectors. The impact does not stop there. When a subset of physicians leave for public service, firms in the private sector face a different population of physicians. For example, if the scheme in Figure 4 is used by the government, then physicians with lower degrees of altruism remain in the private sector. Formally, this is equivalent to a change in the distribution of the physician's altruism parameter. The support of β will have a subset at the top truncated, and the optimal mechanism will have to adjust for this change.

It is beyond the scope of this paper to study fully the equilibrium between the public and private sectors when each offers quantity and compensation schemes. Nevertheless, I speculate that private

firms will be able to pay less information rent when the government hires some physicians. The incentive constraints may become more relaxed because of the smaller physician population.

5 Concluding Remarks

I have presented two models with heterogenous physician behaviors. In the first, some physicians provide health care with good qualities despite the lack of incentives. They work in the public sector. Other physicians behave in strategic ways. When the strategic physicians are allowed to refer patients from the public sector to their private practices, the government may save costs. Welfare increases as a result. I also consider adverse reactions when referrals are possible. They tend to reduce the welfare gain, but public policies that limit the profits to be made by strategic physicians may remedy these adverse effects.

In the second model, physicians possess private information about their own degree of altruism towards patients' benefits as well as patients' valuation of health care. The optimal mechanism offered by competitive firms in the private sector is presented. I then consider how low-powered incentive schemes in the public sector may attract some of the physicians. I illustrate how the public sector may select those physicians with higher degrees of altruism.

In both models I deviate from the conventional wisdom by assuming a variety of behaviors and preferences. I argue that the narrow optimizing behavior framework is inappropriate for the health sector. This is a significant departure. It incorporates a certain degree of realism. I do not in any way suggest that optimizing behaviors are unimportant, but argue that they should be only part of a spectrum of behaviors to be considered.

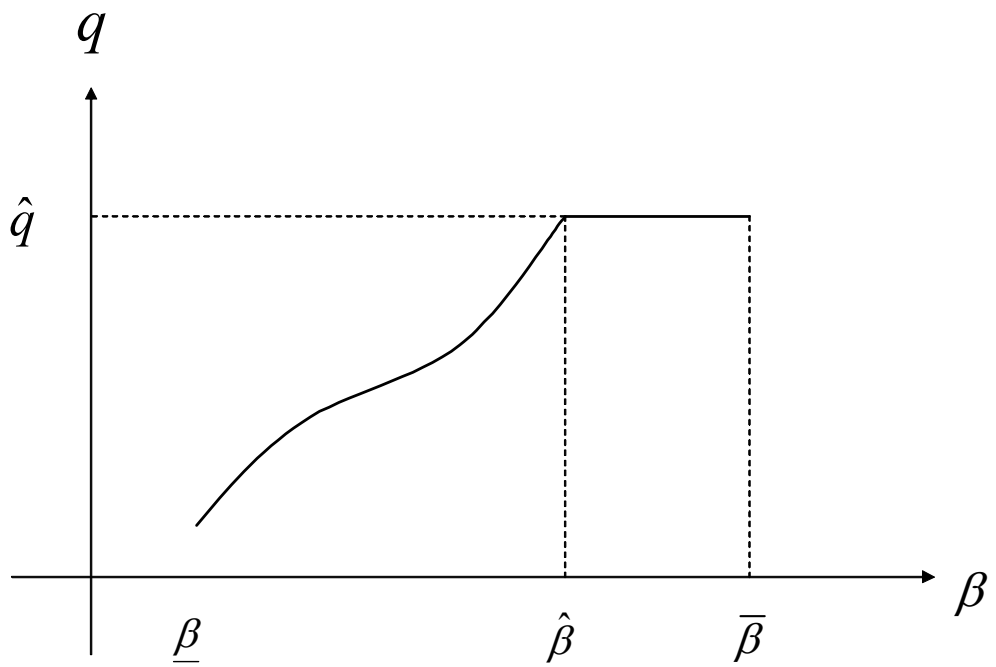

Figure 1: Optimal quantity

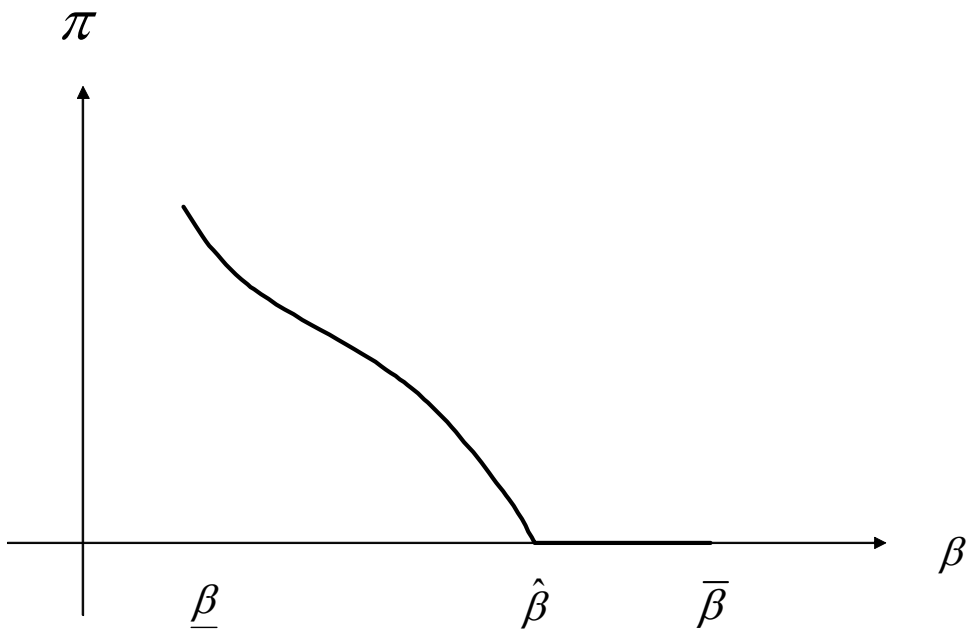

Figure 2: Physician profit

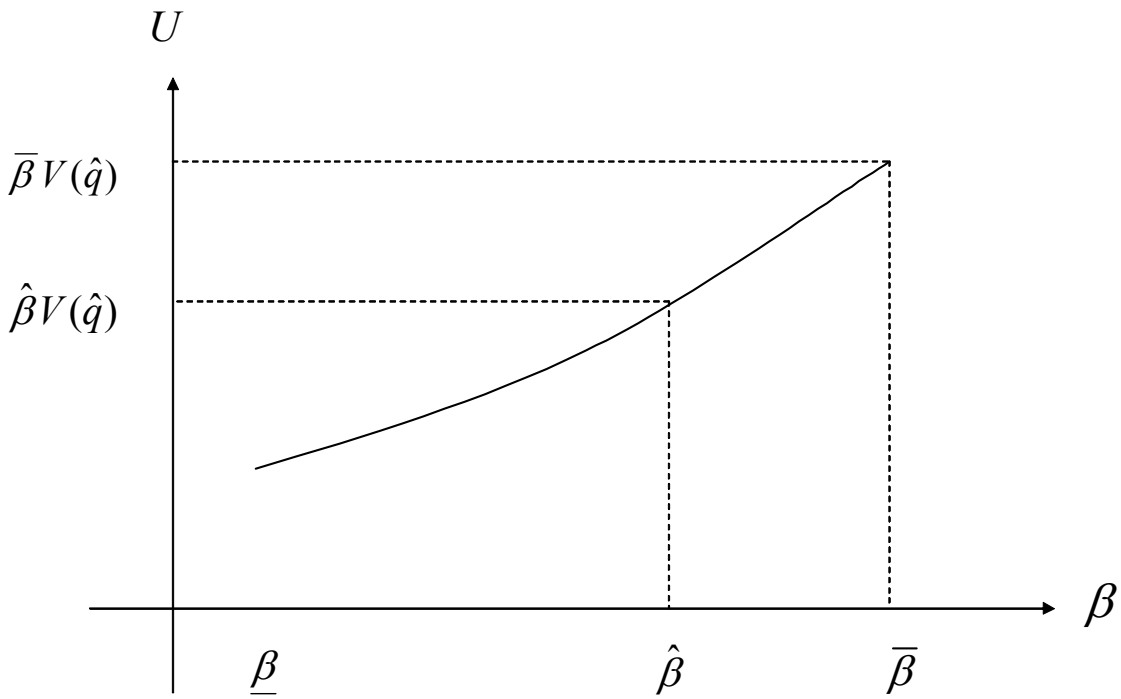

Figure 3: Indirect utility

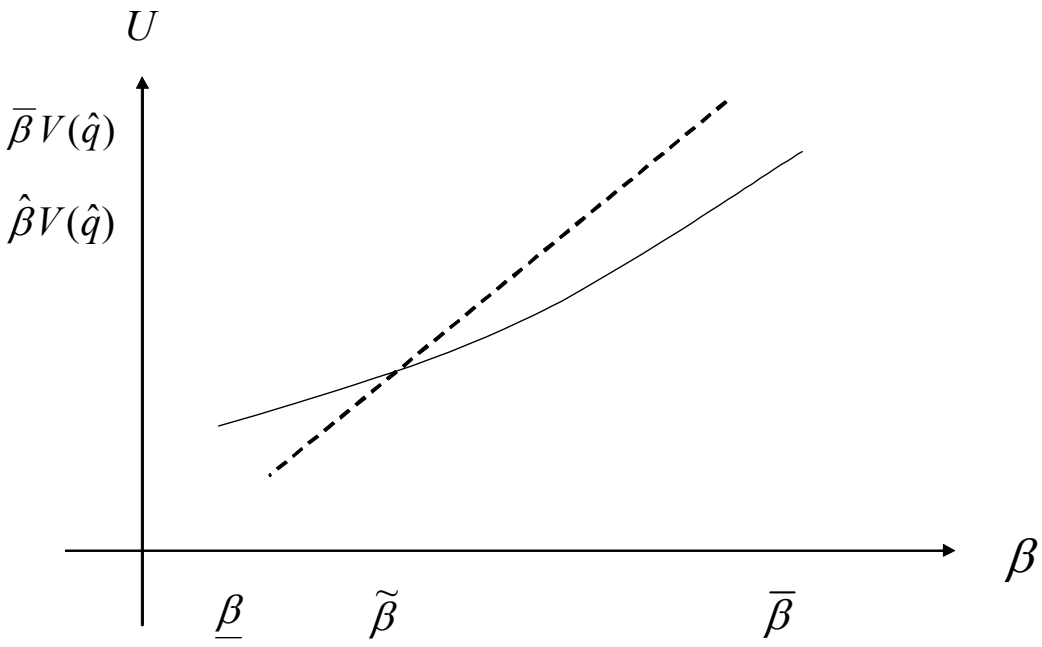

Figure 4: Indirect utilities for different sectors

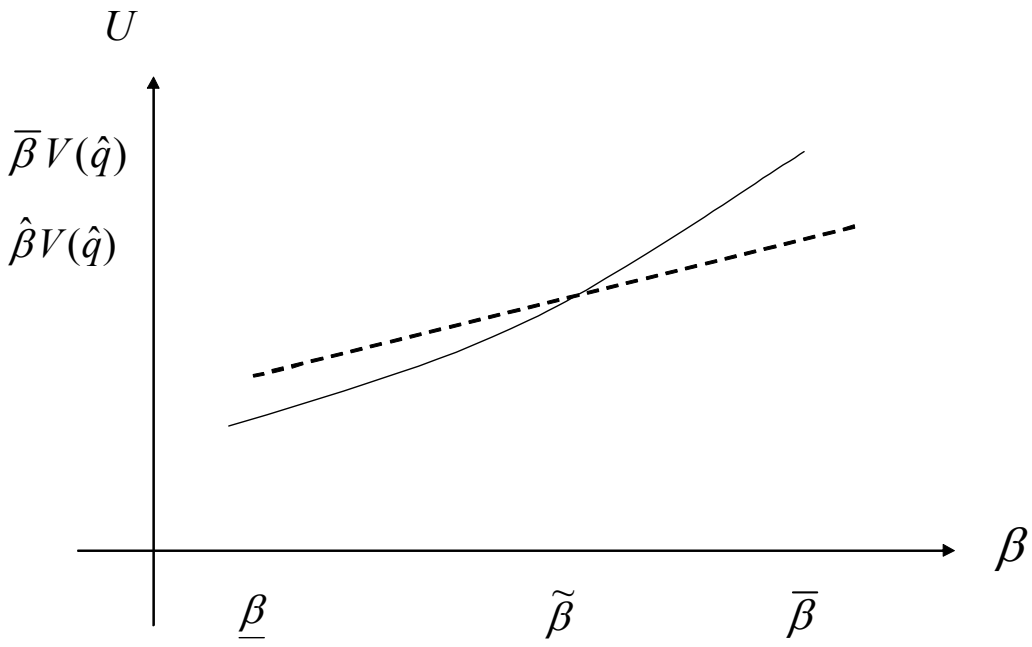

Figure 5: Indirect utility with salary in public sector

References

[1] Alger, I, and Ma, C.A. (2003), "Moral Hazard, Insurance, and Some Collusion," *Journal of Economic Behavior and Organization*, 50, pp. 225-247.

[2] Biglaiser, G. and Ma C.A. (2003), "Moonlighting: Public Service and Private Practice," Boston University working paper.

[3] Brekke, K.A., Kverndokk, S., and Nyborg, K. (2003) "An Economic Model of Moral Motivation," *Journal of Public Economics*, 87, pp 1967-1983.

[4] Barros, P. and Olivella, P. (2002), "Waiting Lists and Patient Selection," CEPR working paper.

[5] Besley, T. and Ghatak, M (2003) "Competition and Incentives with Motivated Agents," London School of Economics working paper.

[6] Besley, T., Hall, J., and Preston, I. (1998), "Private and Public Health Insurance in the UK," *Journal of Public Economics*, 42, pp. 491-497.

[7] Che, Y. (1995), "Revolving Doors and the Optimal Tolerance for Agency Collusion," *Rand Journal of Economics*, 26, pp. 378-397.

[8] Choné, P. and Ma, C.A. (2004) "Asymmetric Information from Physician Agency: Optimal Payment and Healthcare Quantity," Boston University working paper.

[9] Gonzalez, P. (2004), "Should Physicians' Dual Practice be Limited? An Incentive Approach," *Health Economics*, 13, pp 505-524.

[10] Frank, R. (1996), "What Price the Moral High Ground?" *Southern Economic Journal*, 63, pp. 1-17.

[11] Holmstrom, B., and Milgrom, P. (1994), "The Firm as an Incentive Scheme," *American Economic Review*, 84, pp. 972-991.

[12] Iversen, T. (1997), "The Effect of a Private Sector on the Waiting Time in a National Health Service," *Journal of Health Economics*, 16, pp. 381-396.

[13] Jack, W. (2004) "Purchasing Health Care Services from Providers with Unknown Altruism," Georgetown University working paper.

[14] Lewis, T. and Sappington, D. (1989), "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, pp294-313

[15] Ma, C.A. (1994), "Health Care Payment Systems: Cost and Quality Incentives," *Journal of Economics & Management Strategy*, 3, pp. 93-112.

[16] Ma, C.A. (2003), "Public Rationing and Private Cost Incentives," *Journal of Public Economics*, 88, pp. 333-352.

[17] Rabin, M. (1993), "Incorporating Fairness into Game Theory and Economics," American Economic Review, Vol. 83, pp. 1281-1302.

[18] Rickman, N. and McGuire, A. (1999), Regulating Providers' Reimbursement in a Mixed Market for Health Care," *Scottish Journal of Political Economy*, 46, pp. 53-71.