Third Party Purchasing of Health Services: Patient Choice and Agency

Martin Chalkley* Fahad Khalil†

April 13, 2005

Abstract

Health care provision is almost universally characterised by third party purchasing in which the provider of health services is reimbursed by an agency (public or private insurer), rather than the patient. We show how a purchaser can manage the incentives that patient choice gives rise to by its own choice of monitoring arrangement. Even though patients are ignorant about their exact medical conditions and insulated from the costs of health care, they can help alleviate incentive problems due to asymmetric information through the choices that they make about whether to be treated. We show that if patients are responsive to variations in treatment, it can be worthwhile to base payment on the health *outcome* achieved rather than upon the treatment delivered. Outcomes-based payments may also be preferable where services are supplied by not-for-profit agencies who are intrinsically concerned with patient welfare.

JEL classification: I11

Keywords: health contracts, asymmetric information, third-party payers, outcomes research.

*Department of Economic Studies, University of Dundee, Dundee, DD1 4HN, U.K. *m.j.chalkley@dundee.ac.uk* (corresponding author)

†Department of Economics, University of Washington, Seattle, WA 98195-3330. *khalil@u.washington.edu*

1 Introduction

Health care provision is almost universally characterised by third party purchasing in which the provider of health services is reimbursed by an agency (public or private insurer), rather than the patient. In such cases it is imperative that purchasers monitor suppliers directly since patients themselves are largely insulated from the cost of the health care they consume. However, to the extent that quality of service determines the willingness to consume, purchasers can rely on some demand-side discipline imposed by patients. While it is well known that consumers can act as monitors, we show that monitoring schemes differ in their effectiveness in utilizing the disciplining role that patients play. In particular, we show that a health service purchaser's choice of the monitoring arrangement may depend on the strength of patients' demand response to variations in quality of service.

Following Arrow (1963) we consider that the knowledge of an individual patient that is acquired by their physician goes beyond what can be described by the patient's diagnosis and current health state. This asymmetric information is the source of the incentive problem in our model. To induce providers to deliver appropriate treatment, they will have to be given incentives to reveal their information correctly, which necessitates paying them an informational rent. Payments based on the treatments chosen by the health care supplier are commonplace and form the basis of some Diagnosis Related Groups (DRGs) payments in Medicare — see McClellan (1997). Payments based on health outcomes are made possible by the increasing information that is being generated about the effectiveness of different treatments — information that flows from the *outcomes movement* which we briefly review below — and are becoming a reality[1] as third party payers look to provide stronger incentives for health

[1] Examples of this form of payment are reported in the article "Florida Employers Will Offer

care delivery.

In our model treatment-based payments have the advantage that they place a tighter constraint on the supplier's choice of treatment and therefore, *a priori*, would appear to be preferable. However, the analysis indicates that there is an important role for the information that patients have regarding the nature of medical interventions specifically when patients use this information to decide whether to be treated. With patients whose demand is responsive to treatment, a supplier who adjusts treatment so as to elicit higher payment must do so in a way that will cause patients to abstain from being treated. This abstention hurts the supplier and can best be exploited by a purchaser by making payment depend on outcomes. Hence, we conclude that the greater is the demand response on the part of patients the more likely it is for the purchaser to prefer outcome-based payment. A similar argument applies if the supplier is concerned *per se* with the treatment that they offer, perhaps because they are altruistic and value the health gain enjoyed by the patients that they treat. Thus, a purchaser seeking to minimise the cost of health care must chose whether to condition payment upon treatment or on health outcome, paying attention to how these payments make use of patients' demand responses.

Our analysis implies that health outcomes data may be used differently depending upon which medical conditions the health services being purchased are intended to alleviate. In cases of emergency care and life threatening illness like cardiovascular diseases where it is difficult to believe that patients make choices about whether to be treated, payment based on treatment provided is most likely to reduce costs. For elective procedures and less severe illnesses, like ear infections and eczema, it is more likely that purchasers will benefit from conditioning payment on health outcomes.

Incentives to Doctors" in the New York Times 16 November 2001.

This is because responsiveness of patient demand is more likely to be an issue in these cases, and outcome based payments make better use of demand responsiveness. This approach also provides criteria to decide how to channel investment in outcomes research: we establish a cost containment role for outcomes data and show that to the extent that incentive schemes affect the cost of providing health care, additional dollars spent on outcomes research will be *more* effective in types of illness where patient demand is responsive to treatment intensity.

The Outcomes Movement

The outcomes movement, which has also been referred to as the "Third Revolution in Medical Care"[2] (Relman, 1988), has attracted substantial government funding[3], generated intense debate in the medical profession (Epstein, 1990, Naylor, 1995, and Tanenbaum, 1993) and continues to be a major priority in medical research. The movement has two distinct aspects. First, outcomes research considers how an individual's health status, both prior to and following treatment can be measured[4]. This research has emphasized that health status is multi-dimensional and that measures of health status vary from one medical condition to another. Second, the outcomes movement is concerned with understanding the impact of different treatments on im-

[2] According to Relman (1988), the first two 'revolutions' are the expansion of health services (in the 1940s and 1950s) and the movement to contain costs (in the 1980s).

[3] In 1989 the federal Agency for Health Care Policy Research was set up in the US, explicitly committed to conduct and disseminate outcomes research. In the British National Health Service outcomes research is central to the process of Clinical Audit (for which health providers receive a specific budget) and Evidence Based Medicine (Whynes, 1996).

[4] Published outcomes research provides many indications of this focus. For example, papers from the Centre for Health Quality, Outcomes and Economic Research (http://dcc2.bumc.bu.edu/chqoer/about.htm) contain the following: 'Our objective was to develop a patient-based measure of the severity of osteoarthritis of the knee focusing on symptomatology, that may be used in conjunction with measures of health-related quality of life in monitoring the health status of outpatients.' and 'We developed a symptom-based measure of severity for chronic lung disease(CLD) that can be readily administered in ambulatory care settings and be used to supplement general health-related quality of life (HRQoL) assessments and pathophysiologic indicators in research and clinical care.'

provements in health status[5]. This activity has centred on the collection and analysis of large data sets where information on changes in health status, together with information on treatments undertaken is recorded and analysed. Proponents of this research see it as establishing the proper basis for measuring medical output and, hence, claim that it will "bring order and predictability" (Ellwood, 1988) to health care systems. Such claims have a basis in economic analysis because if the outcomes research program succeeds in isolating the medically effective treatments for a range of conditions, it will provide purchasers of health services with additional information that can be used to better specify the reimbursement of health care suppliers. This is the reason for our focus on the outcomes movement and our subsequent analysis of payment mechanisms that might be derived from it.

Outcome measures, in the form of infection rates, morbidity rates and related measures, are now routinely collated. The QALY — a measure of the impact of health care interventions on patient welfare — is commonly used when discussing the allocation of health care resources[6]. For example, Dranove (2000) reports, "...interventions that cost more than $60,000 per QALY are often deemed cost-ineffective". He also points out that , "this threshold is considerably lower than the value of a year of life identified by economic studies". As noted above the outcomes movement has emphasized the multi-dimensional nature of many health care interventions so that a single summary measure is questionable. Nevertheless it is possible that the concept of the QALY may be refined, or extended to many dimensions, such that there is broad agreement that some such measure is an appropriate indication of the effec-

[5] For example, Tanenbaum (1993) comments thus on the Medical Treatment Effectiveness Program: 'This program differs from earlier efforts in its focus on medical effectiveness and its sponsorship of large-scale statistical studies of both common and alternative treatments for specific conditions. It is a part of what has been called the "health-outcomes strategy"'.

[6] See Lu (1999).

tiveness of a particular health care intervention. In which case it would correspond to the outcome measure that we consider below.

Related papers

Central to our analysis is the concept of patients responding to the treatments they are offered by choosing where, or whether, to be treated. The idea that patients who have a long term relationship with a physician, or who receive information from friends, can be expected to choose where to be treated — or in the case of elective treatments whether to be treated at all — has been an important aspect of the analysis of health contracts. Previously, such a *demand response* on the part of patients has been identified as an important incentive instrument in mitigating the effects of *moral hazards* which are perceived to lead to excessive cost and compromised quality of care. Discussion has, in particular, focused on how payment based only on the number of patients treated (*price-quantity schedules*) can be used to provide appropriate incentives – see for example Ma and McGuire (1997), Ma (1994), McGuire (2000) and Chalkley and Malcomson (1998a, 1998b, 2000). In the US price quantity schedules are inherent in the prospective payment system used by Medicare from 1983 and in the UK health purchasers have been encouraged to adopt similar arrangements since the early 1990s. In contrast, we find that demand effects also help to align incentives by reducing provider rents due to asymmetric information. But, more significantly, we show that outcomes-based payment is more effective than treatment-based payments[7]

[7] The seminal work by Maskin and Riley (1985) establishes the theoretical framework concerning the choice of monitoring instruments in contracts with asymmetric information that is the foundation of this paper. They show that input-based taxes are more efficient than output-based taxes. More recently Khalil and Lawarrée (1995) have shown that output-based schemes can be superior when the agent receives a transfer. Lewis and Sappington (1995) and Bontems and Bourgeon (2000) are other recent contributions to this literature. In contrast to these authors, we show how the choice of instruments determines the effectiveness of the monitoring role played by patients, which in turn may determine the choice of the monitoring instrument itself. The central insight we offer is that, with asymmetric information, the choice of the payment scheme may depend on the responsiveness of demand by patients.

in utilizing patients' demand response.

Asymmetric information regarding patient types of the kind that we consider here has featured in the work of Dranove (1987), Allen and Gertler (1991), Ma (1994), and Ellis and McGuire (1986). In this literature the focus is upon the effect of using a single payment to cover patients of different types with a view to examining the incentives generated by prospective payment systems as an alternative to cost reimbursement in health care markets. In contrast to these papers, the present paper considers how a purchaser may best fine tune payments so as to ensure that different types of patients receive the kind of treatment that is efficient for them. The alternatives with a single payment covering multiple types are either that some patients are not treated (or *dumped* in the terminology of Ma, 1994) or that suppliers earn excessive profits by choosing to treat easier patients, a process referred to as *cream-skimming* (see Barros, 2003). Without the information that it is provided by outcomes research, a purchaser has little option but to choose between dumping or cream skimming. Hence, our approach is predicated on purchasers having more detailed information and considers how that information might best be used. Lewis and Sappington (1999) adopt an approach that is similar to the one pursued in this paper to address a different question, that of how information *acquisition* by suppliers affects the form of the optimal contract.

The organization of the paper is as follows. A model of health service provision with private information is presented in section 2. The optimal contract under full information is derived in section 3. In sections 4 and 5, we study the case of treatment-based and outcome-based payments, and compare the two in section 6, where the main result of the paper is presented. We consider supplier altruism in section 7 and discuss the results and their implications for health care policy in section 8.

2 The Model

We consider a purchaser who contracts with a single[8] health care supplier, which for convenience we refer to as the *hospital*, in order to ensure the provision of treatment of patients with a particular medical condition and a given health status prior to treatment.[9] Patients may be one of two *types,* which we denote by $\phi \in \{\phi_g, \phi_b\}$ where $\phi_g > \phi_b$.Whilst we restrict attention to a model with two types for expositional simplicity, we show in the appendix how our results would easily generalize to a model with multiple types. A patient of *good* type ϕ_g is assumed to respond well to treatment and be cheap to treat whereas the opposite is assumed for *bad* type ϕ_b. The purchaser, hospital and individuals all share the same *ex ante* assessment of the probability of a patient's type, where $\Pr[\phi = \phi_g] = \pi$ and $\Pr[\phi = \phi_b] = 1 - \pi$. An individual's type depends on the precise nature of their condition as can be determined solely by a physician, and is therefore unknown to either the purchaser or the individual. A patient's type is, however, discovered by the hospital at the onset of treatment and, hence, there is asymmetric information.

For each patient, the hospital determines an intensity of treatment[10], which we denote by x. A type-ϕ patient given a treatment of intensity x will have a gain in health status[11] of $h(x, \phi)$ increasing in x and ϕ. The purchaser attaches a monetary

[8] We treat the hospital as a monopoly supplier and thus abstract from difficulties that arise when several hospital suppliers compete for a contract.

[9] For the purposes of exposition we consider a medical condition of a given severity. In practice, a contract could be written to cover both a *diagnosis* and *severity* of condition. Hence one diagnosis could give rise to many conditions being contracted for.

[10] For convenience we consider treatment intensity as a scalar quantity. Intensity can also be thought of as an index of a multidimensional vector that characterizes a particular medical intervention. Provided that the different dimensions of treatment occur in fixed proportions the analysis is unaffected. When there are truly multi-dimensional aspects to a hospital's decisions then new issues arise of the kind discussed in Chalkley and Malcomson (1998a).

[11] We treat h as a scalar and deterministic purely for convenience. It is straightforward to allow for a vector of characteristics representing an individual's health status and for h to be a random variable with density $f(h\,|x, \phi)$.

value $v(h)$ to the gain in health status h, and the purchaser's benefit is given by $b(x, \phi) \equiv v(h(x, \phi))$ We assume that $b(x, \phi)$ is increasing and concave in x, increasing in ϕ and such that the marginal benefit of treatment is non-decreasing in ϕ. These assumptions ensure that the purchaser's expected benefit function is well behaved and meets the requirement that treatment is more effective for good types. We assume that the cost of treating a patient of type ϕ with intensity x can be written as $c(x, \phi)$ and assume that $c(.)$ is increasing and convex in x, decreasing in ϕ and that the marginal cost of treatment is non-increasing in ϕ. Again, these assumptions ensure that costs are well-behaved and meet the requirement of treatment being less costly for good types.[12] We assume that $c(x, \phi)$, which we take to be the true economic cost of treatment, cannot be observed by the purchaser. This is consistent with, for example, financial costs being observable but there being elements of cost that are not reported and are private information to the hospital. We also assume the necessary Inada conditions so that we obtain positive but bounded values for choice variables at the optimum.

Our model of demand response by patients follows that of Ma and McGuire (1997) and McGuire(2000). We assume that long-term relationships between hospitals and patients or information from friends allows individuals to form an assessment of the intensity of treatment that is on offer from the hospital. Since more intense treatments increase health status, patients will choose where to be treated, or in the case of elective procedures whether to be treated at all, according to this assessment of treatment intensity. We therefore assume that demand is a function of the intensity of treatment that individuals expect[13] and that individuals prior to their own treatment

[12] According to our assumptions, we have a clear ranking of types according to good and bad prospects. See Lewis and Sappington (1989) and Bontems and Bourgeon (2000) for a discussion of the effects of countervailing incentives in a related models.

[13] This, in common with much of the literature on health contracts, presumes that the health treat-

have an unbiased signal of the average intensity of treatment they will receive if treated. Hence, if good types are treated with intensity x_g and bad types with intensity x_b each patient anticipates that they will receive treatment of expected intensity $\bar{x} = \pi x_g + (1-\pi)x_b$. We suppose that the total number of patients who wish to be treated[14] is an increasing function of expected intensity $\bar{x}$, and so denote total *expected* demand[15] for treatment as $n(\overline{x}(x_g, x_b))$, with $n'(\bar{x}) \geq 0$. The special case in which $n' = 0$ corresponds to patients who are either ignorant of the treatment intensity that they will receive or cannot respond to changes in expected intensity, e.g., because of medical emergencies. Since patients do not know their type, $n(.)$ is independent of ϕ.

We assume initially that the hospital operates "for profit" but consider in section 7 the implications of the hospital having a concern for its patients. Since the hospital observes patients' types it can choose treatments conditional on type. A *treatment policy* for the hospital consists of a type contingent treatment intensity for each patient which we write as $\{x_g, x_b\}$. The hospital's treatment policy determines both the revenue and cost of each patient treated and expected demand.

Payment from the purchaser to the hospital depends on the patient's type as reported by the hospital. Under what we call *treatment-based payment,* the purchaser verifies that patients receive the treatment according to the types claimed. Under what we call *outcome-based payment,* the purchaser verifies that patients' health-gains

ments being considered have the attributes of *search* goods. Our main results would be preserved if we assumed that demand was determined by *expected* benefits.

[14]This reduced form for demand can be rationalised by assuming that patients' utility depends on expected treatment and improvement in health status, and belief about type and that they seek treatment if utility is above a reservation value. Using a distribution of reservation values, there would be demand that is a function of treatment intensity, belief about responsiveness of health status to treatment and belief about patient type.

[15]Random demand together with the requirement to ensure that patients are not dumped, which is formalized subsequently in individual rationality constraints, will make it infeasible to base penalties on the number of patients treated.

(i.e., the outcome) corresponds to the types claimed. We assume that the purchaser can impose sufficient penalties such that the hospital will want to ensure that what the purchaser observes is consistent with its claim. Hence, if a type ϕ_g patient is claimed by the hospital to be type ϕ_b they will receive a treatment of intensity x_{bt} under treatment-based payment, whereas under outcomes-based payment, they will receive treatment with an intensity that results in a health-gain equal to that of a type ϕ_b patient. The hospital may thus misrepresent a patient's type if it is in its interest to do so.[16]

To isolate the incentive effects of each payment scheme we further assume that payment is based on *either* treatment alone *or* on outcome alone but that the costs of implementing payment are the same in each case. This is equivalent to assuming that there is a fixed cost of setting up each monitoring system such that the purchaser must choose whether to verify a hospital's claims in respect of either treatment or outcomes because verifying both, and thus incurring two fixed costs, is prohibitively expensive. In practice the fixed cost of establishing a system to verify outcomes may differ from the fixed cost of establishing a system to verify treatments, but the implications are obvious. If both treatment and outcome could be observed in our model, it is equivalent to observing a patient's type, and the first best (net of verification cost) can be implemented. In practice the purchaser-supplier relationship is complex so that purchasers would not eliminate asymmetry of information even if they pursue elements of both treatment and outcome verification. In practice purchasers may, therefore, want to incorporate both in their reimbursement systems. Our analysis acts as a guide to the relative merits of each payment system by focusing on the

[16] See Alger and Ma (2003) for references to the recent literature where agent's are assumed to be "honest" with an exogenous probability. They show that the second best contract, which we rely on, is optimal as long as this probability is not too large

incentive effects of one instrument at a time.

The timing of events that we assume is as follows. To start with, everyone has symmetric belief about patient types. The purchaser designs a contract under which the hospital will be rewarded for each person treated subject either to verification by monitoring of treatments given or of outcomes achieved. Prior to receiving treatment, individuals receive a signal (the average intensity of treatment) and decide on the basis of that signal whether or not to be treated. Those deciding to be treated then go to the hospital to receive treatment. Only the hospital learns the type of the patient prior to offering treatment. Following treatment by the hospital, the purchaser assesses either the intensity of treatment that a patient undergoes or the outcome of treatment and the hospital is then paid according to its contract with the purchaser.

3 Full Information

If there were full information on patient types, the purchaser could condition payment directly on a patient's type. Then the purchaser's objective function[17] is

$$n(\bar{x})\left[\pi\left(b(x_g,\phi_g)-p_g\right)+(1-\pi)\left(b(x_b,\phi_b)-p_b\right)\right], \tag{1}$$

where p_g, p_b denote transfers paid by the purchaser for, respectively, good and bad type patients. The purchaser needs to ensure that the hospital is willing to provide the necessary treatment to each type and so must ensure that the hospital makes a non-negative return on each type of patient to avoid 'dumping'[18]. We therefore

[17] For expositional simplicity we assume that the purchaser is concerned with the health gains of patients but does not attach any weight to the hospital's profit. Our qualitative results are preserved if the "cost of public funds" approach, as in Laffont and Tirole (1993), is used instead.

[18] See Lewis and Sappington (1999) and Ma (1994) for more on dumping and cream skimming. McClellan (1997) provides evidence that diagnostic related groups (DRG) are often defined to accommodate exceptional cases with high treatment cost in order to avoid dumping.

assume that the optimal contract will satisfy the individual rationality constraints:

$$\begin{aligned} p_g - c(x_g, \phi_g) &\geq 0, && (IR_g) \\ p_b - c(x_b, \phi_b) &\geq 0. && (IR_b) \end{aligned}$$

The purchaser's problem is to maximize the objective function (1) subject to IR_g and IR_b. Since in (1) any transfer that the purchaser makes to the hospital subtracts from its welfare, the two constraints are binding, and we can substitute for p_g and p_b and solve for an unconstrained optimum[19]. To simplify notation, we define expected surplus per patient as

$$S(x_g, x_b) \stackrel{\text{def}}{=} \pi \left(b(x_g, \phi_g) - c(x_g, \phi_g)\right) + (1 - \pi) \left(b(x_b, \phi_b) - c(x_b, \phi_b)\right) \tag{2}$$

and write the purchaser's objective function as

$$W(x_g, x_b) \stackrel{\text{def}}{=} n(\bar{x}) S(x_g, x_b). \tag{3}$$

Denoting partial derivatives by subscripts and the derivative of demand with respect to average intensity by n', the first best treatment intensities are the solutions to the following first order conditions:

$$W_g(x_g^*, x_b^*) = n'(\bar{x}^*) S(x_g^*, x_b^*) + n(\bar{x}^*) \left[b_x(x_g^*, \phi_g) - c_x(x_g^*, \phi_g)\right] = 0, \tag{4}$$

$$W_b(x_g^*, x_b^*) = n'(\bar{x}^*) S(x_g^*, x_b^*) + n(\bar{x}^*) \left[b_x(x_b^*, \phi_b) - c_x(x_b^*, \phi_b)\right] = 0, \tag{5}$$

[19] The program defined by maximizing (1) subject to IR_g and IR_b is assumed to be well behaved with a unique optimum. In the absence of the function $n(.)$, concavity of $b(.)$ and convexity of $c(.)$ would ensure a well-behaved program. However, an increase in treatment increases demand and if this effect is strong enough, we could have an unbounded solution. Therefore, we are assuming that the properties of $n(.)$ do not invalidate the convexity of the program implied by standard assumptions on the functions $b(.)$ and $c(.)$.

where $\bar{x}^*$ denotes $\bar{x}$ evaluated at first best treatment intensities. These two conditions illustrate the effects of demand on optimal treatment intensities. Since demand is increasing in treatment intensities, the first-best intensity of treatment for each type is extended beyond that which makes marginal benefit equal marginal cost for a given patient. The extent of this excess of treatment (over that which would prevail in the absence of patients responding to treatment intensity by demanding treatment) is captured by the terms involving n'. The first order conditions also make clear, given our assumptions on $b(.)$ and $c(.)$, that a good type will receive more intensive treatment than a bad type in the first best, i.e. $x_g^* > x_b^*$.

For future reference it is useful to note that, with $n'' \leq 0$, the cross partial $W_{gb}(x_g^*, x_b^*)$ is negative. This follows because an increase in the treatment intensity offered to bad types decreases average surplus $S(.)$ and reduces the value on the margin from treating good types, since the net marginal benefit is negative $([b_x(x_g^*, \phi_g) - c_x(x_g^*, \phi_g)] < 0,)$ at the first best treatment. By symmetry, $W_{bg}(x_g^*, x_b^*) < 0$ and the same argument applies.

4 Treatment-based payment

We now consider asymmetric information where the purchaser chooses to observe the treatments provided but does not observe the patient's type nor the health status improvement due to treatment. The payment to the hospital is a function of treatment, and the contract offered to the hospital is $\{p_{gt,}p_{bt}, x_{gt,}x_{bt}\}$. If the purchaser were to offer the hospital the first best payments $p_g = c(x_g^*, \phi_g), p_b = c(x_b^*, \phi_b)$ the hospital would have an incentive to misrepresent a type ϕ_g as a type ϕ_b because whilst it will receive a lower price, it will incur a substantially lower cost and thus earn a rent of $\left[p_b - c\left(x_b^*, \phi_g\right) > 0\right]$. Anticipating these incentives, the optimal contract is the solution to the purchaser's problem P_t written below.

$$\max \left[n(\bar{x}) \left\{\pi \left(b(x_{gt}, \phi_g) - p_{gt}\right) + (1-\pi)\left(b(x_{bt}, \phi_b) - p_{bt}\right)\right\}\right]$$

$$\text{subject to} \tag{6}$$

$$n(\bar{x}(x_{gt}, x_{bt})) \left[p_{gt} - c(x_{gt}, \phi_g)\right] \geq n(\bar{x}(x_{bt}, x_{bt})) \left[p_{bt} - c(x_{bt}, \phi_g)\right], \tag{7}$$

$$p_{bt} - c(x_{bt}, \phi_b) \geq 0. \tag{8}$$

The first constraint is the incentive compatibility constraint for a good type. It ensures that the hospital cannot gain by treating good type patients as if they were bad. The second constraint is the individual rationality constraint applying to bad type patients. There are two other constraints: the incentive constraint for a bad type, and individual rationality constraint for a good type. Neither are included in the definition of P_t because, as is typical in models of this type, they are not binding in equilibrium. It can be easily verified that the optimal contract satisfies the omitted constraints as inequalities.

The two constraints (7) and (8) will hold with equality since the payment to the hospital can otherwise be lowered to the purchaser's benefit. The left hand side of (7) measures the total rent that the hospital will earn under a treatment-based payment. Substituting from (8) into the right hand side of (7) and imposing equality, we can obtain an expression for this rent[20] as:

$$\begin{aligned} R_t(x_{bt}) &= n(\bar{x}(x_{bt}, x_{bt})) \left[c(x_{bt}, \phi_b) - c(x_{bt}, \phi_g)\right], \\ &> 0. \end{aligned} \tag{9}$$

[20] The inequality requires $x_{bt} > 0$, but that will be true in equilibrium.

In the solution[21] of the second best problem the purchaser must pay the hospital transfers which exceed the cost of treatment in order to ensure that the hospital has an incentive to offer appropriate treatments to each type of patient. The net value of these transfers is given by (9). The following proposition makes precise which patients a hospital will earn rents on, and the implications of these rents for treatment intensities. In the proposition and subsequently, we use $\tilde{\cdot}$ to denote second best treatment intensities.

Proposition 1 *The hospital receives rent from treating good type patients but none for treating bad types. There is under-provision of treatment intensity for bad types i.e. $\tilde{x}_{bt} < x_b^*$, and good types always receive higher intensity than bad types i.e. $\tilde{x}_{gt} > \tilde{x}_{bt}$. It is ambiguous whether there will be over or under-provision of treatment intensity (relative to the first-best) for good types.*

Proof. In appendix

The nature of the distortions in treatment intensities can be understood by examining the effect of x_{bt} on the expected rent $R(x_{bt})$. Since the cost differential $\left[c(x_{bt}, \phi_b) - c(x_{bt}, \phi_g)\right]$ increases with x_{bt}, the expected rent increases with x_{bt}. Therefore, the purchaser will want to lower x_{bt} from the first best amount in order to reduce expected rent. Under-provision of treatment to the bad type is a standard result, and if it were not for a demand effect, we would also have the standard result that x_g is first best. However, changes in x_b affect the choice of x_g via $n(\cdot)$, and this effect is ambiguous in general. Interestingly, a reduction in x_b from x_b^* can result in an increase in x_{gt} above the first best level even though $R_t(.)$ is independent of x_{gt}.

[21] To ensure a well behaved program under asymmetric information, it is sufficient to assume that $n''(.)$ is small, or not too negative, in addition to assumptions made under full information. This additional assumption makes rent convex in x_b.

By assumption patients are insulated from the cost of their treatment and, hence, there is a direct link between treatment intensity and patient welfare. Relative to the first best, type ϕ_b patients are worse-off under second best treatment-based payment, but type ϕ_g patients may be better-off or worse-off. Consider an example where π is very small. In that case, $\tilde{x}_{bt}$ is close to but smaller than x_b^*, and we already know from the analysis of the full-information problem that $W_{gb}(x_g^*, x_b^*) < 0$. Since $W_{gg} < 0$ by assumption, for $\tilde{x}_{bt}$ close to but smaller than x_b^*, we must have $\tilde{x}_{gt} > x_g^*$. In general, however, $W_{gb}(.)$ is ambiguous, and there may be under-provision of intensity, and hence lower welfare for the good types.

Since $\tilde{x}_{gt} > \tilde{x}_{bt}$, demand responsiveness acts as a disciplining device because, as can be seen from expression (9), as the hospital adjusts treatment so as to elicit higher payment it must do so in a way that will cause patients to abstain from being treated, which lowers total rent.

5 Outcome-based payment

If the purchaser wishes to base payments on observed outcomes, we assume that it does not learn patient type nor the treatment provided. Under outcome-based payment, the purchaser chooses the contract $\{p_{go,}p_{bo}, h_{go,}h_{bo}\}$, where the provider is paid p_{ko} when the observed health gain to a patient is h_{ko}, and $h_{ko} = (x_{ko}, \phi_k)$ for $k = b, g$. Given this contract, the provider chooses treatment levels for each patient. Outcome-based payment imposes a constraint on the treatments that the hospital must give if it is to misreport patient types but satisfy *ex post* verification. Again, the incentive is to misrepresent a type-ϕ_g as a type-ϕ_b for both the reason given above and additionally because in misrepresenting such a patient under outcome monitoring, the hospital can further economise on treatment and, hence, cost. This can be seen as follows: if the hospital wants to misrepresent a good type as bad, it must provide the type-ϕ_g with an intensity of treatment such that the expected health gain is the same as if the patient was type-ϕ_b. Specifically, this intensity of treatment is

$$\hat{x}_{bo} = \hat{x}(x_{bo}, \phi_g, \phi_b), \tag{10}$$

defined by

$$h(x_{bo}, \phi_b) \stackrel{\text{def}}{=} h(\hat{x}_{bo}, \phi_g). \tag{11}$$

In other words, a type-ϕ_g needs to be given a treatment $\hat{x}_{bo} < x_{bo}$ to generate an expected health gain equal to the level relevant for a type-ϕ_b patient. This is illustrated in Figure 1.

Even though treatment levels are not observable, the purchaser will anticipate the incentive of the provider and choose the contract to solve the problem P_o given by.

$$\max \left[n(\bar{x}) \left\{ \pi \left(b(x_{go}, \phi_g) - p_{go} \right) + (1 - \pi) \left(b(x_{bo}, \phi_b) - p_{bo} \right) \right\} \right]$$

subject to

$$n(\bar{x}(x_{go}, x_{bo})) \left[p_{go} - c(x_{go}, \phi_g) \right] \geq n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[p_{bo} - c(\hat{x}_{bo}, \phi_g) \right], \tag{12}$$

$$p_{bo} - c(x_{bo}, \phi_b) \geq 0, \tag{13}$$

As in the previous section, the first constraint is the incentive compatibility constraint for a good type, and the second is the individual rationality constraint when the type is bad. Also as previously, the other incentive constraint and individual rationality constraint, which are satisfied as inequalities in equilibrium, are omitted.

The two constraints must be binding, otherwise the purchaser can lower p_{go} and p_{bo} to his benefit. Applying the same method as used in the section on treatment-based payment the rent the hospital will earn under a second best outcome-based contract can be written

$$
\begin{aligned}
R_o(x_{bo}) &= n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[c(x_{bo}, \phi_b) - c(\hat{x}_{bo}, \phi_g) \right], \qquad (14) \\
&> 0.
\end{aligned}
$$

.

Proposition 2 *The hospital receives rent from treating good type patients but none for treating bad types. There is under-provision of treatment intensity for bad types i.e.* $\tilde{x}_{bo} < x_b^*$*, and good types always receive higher intensity than bad types i.e.* $\tilde{x}_{go} > \tilde{x}_{bo}$*. It is ambiguous whether there will be over or under-provision of treatment intensity (relative to the first-best) for good types.*

Proof. In appendix.

The explanation of this proposition follows that given in the case of treatment-based payment.

6 Outcome-based payment makes better use of patients' demand response

Under both the payment systems discussed above, the purchaser pays the hospital a rent on account of asymmetric information. The different payment schemes have different implications for the treatments that patients receive but there is always under-treatment for bad types relative to the first best and may be either under- or over-treatment of good types. The precise extent of the distortions that arise from information asymmetry will vary according to functional form and parameters. Here we are concerned with the relative cost to the purchaser of adopting treatment or outcome-based payments as that is measured in the rent required to implement a particular set of treatment intensities. Specifically we are concerned with knowing whether one form of payment system dominates the other. The following proposition provides the details.

Proposition 3 *For any pair of treatment intensities* (x_g, x_b) *that are to be implemented under treatment-based payment, outcome-based payment will implement those intensities at a lower overall cost to the purchaser if:* $n(\bar{x}(x_b, x_b))\left[c(x_b, \phi_b) - c(x_b, \phi_g)\right] > n(\bar{x}(\hat{x}_b, x_b))\left[c(x_b, \phi_b) - c(\hat{x}_b, \phi_g)\right]$*, where* $\hat{x}_b$ *is obtained by substituting* x_b *for* x_{bo} *in (10).*

Proof. Consider the pair of treatment intensities (x_g, x_b) that are to be implemented under treatment-based payment, and define by the pair (p_{gt}, p_{bt}) the minimum transfers needed to implement them. Using the constraints (8) and (7), we have

$$\begin{aligned} p_{bt} &= c(x_b, \phi_b), && (15) \\ p_{gt} &= c(x_g, \phi_g) + \frac{n(\bar{x}(x_b, x_b))}{n(\bar{x}(x_g, x_b))}\left[c(x_b, \phi_b) - c(x_b, \phi_g)\right]. && (16) \end{aligned}$$

We define by (p_{go}, p_{bo}) the pair of (minimum) transfers that implement the pair of intensities (x_g, x_b) under outcome based payment. They satisfy

$$\begin{aligned} p_{bo} &= c(x_b, \phi_b), && (17) \\ p_{go} &= c(x_g, \phi_g) + \frac{n(\bar{x}(\hat{x}_b, x_b))}{n(\bar{x}(x_g, x_b))}\left[c(x_b, \phi_b) - c(\hat{x}_b, \phi_g)\right]. && (18) \end{aligned}$$

Since $p_{bt} = p_{bo}$, we only need to show that $p_{gt} > p_{go}$, which is equivalent to the condition

$$n(\bar{x}(x_b, x_b)) \left[c(x_b, \phi_b) - c(x_b, \phi_g)\right] > n(\bar{x}(\hat{x}_b, x_b)) \left[c(x_b, \phi_b) - c(\hat{x}_b, \phi_g)\right]. \tag{19}$$

■

The result follows from a comparison of rents under the two payment schemes. Rents in both treatment and outcome-based cases have similar features. In both cases the cost differential of providing treatment to a good type and receiving reimbursement as if the type was bad, and the drop in demand due to lower treatment intensity figure as determinants of rent. There are, however, significant differences. Since with outcome-based payment, x_b is larger than $\hat{x}_b(x_b, \phi_b, \phi_g)$, the cost differential is larger in this case, but the drop in demand is also larger. These differences will be greater the larger is the discrepancy between x_b and $\widehat{x}_b$ and the greater is the discrepancy in the marginal cost of treatment intensity. Figure 2 shows that rent per person under outcome-based payment, shown as $\frac{R_o}{n}$, is larger than that under treatment-based payment, shown as $\frac{R_t}{n}$. Whilst the per person rent is higher under outcome-based payment the number of patients determining total rent is lower by an extent that depends upon demand responsiveness. Hence greater demand responsiveness reduces rent under outcome-based payment relative to treatment-based payment. Intuitively, if demand responsiveness is strong enough, the purchaser will prefer an outcome-based system over a treatment-based one.

The condition in Proposition 3 depends on three things. The extent of the difference between the outcome of treatment to good and bad types determines the extent to which x_b is greater than $\hat{x}_b$. The responsiveness of demand to variations in average treatment intensity determines, for any given difference between $\hat{x}_b$ and x_b, the extent to which $n(\bar{x}(x_b, x_b))$ is greater than $n(\bar{x}(\hat{x}_b, x_b))$. Finally, the curvature of the good type's cost function determines the magnitude of

$$c(x_b, \phi_b) - c(x_b, \phi_g) < c(x_b, \phi_b) - c(\hat{x}_b, \phi_g).$$

The interaction between these three effects is complex. However, when there is a negligible impact of variations in treatment intensity of good types on the cost associated with their treatment, the condition is almost certainly satisfied. We therefore, have as a limiting case:

Corollary 4 *If $c_x(x, \phi_g) = 0$ (for all x) outcome-based payment results in a lower overall cost to the purchaser than treatment-based payment.*

Proof. Follows from (19) by using $x_b > \hat{x}_b$ and $c_x(x, \phi_g) = 0$ (for all x). ■

It is also possible to consider circumstances under which the condition will not be satisfied. The most obvious case being where average treatment intensity does not impact on demand. Therefore,

Corollary 5 *If $n'(.) \equiv 0$ and $c_x(x, \phi_g) > 0$, the cost to the purchaser of treatment-based payment is lower than the cost of outcome-based payment.*

Proof. Follows from (19) by setting $n(\bar{x}(x_b, x_b) = n(\bar{x}(\hat{x}_b, x_b))$.

These results indicate in what ways differences in demand and cost will influence the optimal form of payment by purchasers. Different medical conditions vary both in the extent to which patients are aware of the treatment they will receive and the extent to which patients can be expected to respond to variations in treatment, and will be characterized by different marginal costs of treatment intensity. Where patients perceive and respond to intensity of treatment, as is likely to be the case for elective procedures, the analysis suggests that even if purchasers are equipped with a detailed knowledge of what treatments are effective for each medical condition they may still wish to monitor the effectiveness of treatment for themselves and condition payment on what they observe. For medical conditions such as, for example, emergency treatments, where there is little exercise of choice by patients (and, hence $n'(.) \approx 0$) the analysis indicates that purchasers will do better by conditioning payment on the treatments given.

7 Altruism

The analysis above assumes that the hospital maximizes profit whereas there is empirical evidence, such as that presented by Dranove and White (1994), which indicates hospitals may be motivated by a concern for the patients that they treat. In a not-for-profit hospital it has been argued that treatment intensity may be of intrinsic concern to the supplier – see, for example, Newhouse (1970). We capture the potential concern that a hospital might have for its patients by considering the hospital as having an altruistic component to its objective function of $A(x, \phi) = a(h(x, \phi))$, where $a(h)$ is the hospital's valuation of h. The main result we obtain is that altruism will bias the purchaser towards outcomes-based payments.

The formulation of altruism that we use presumes that there is a limit on the ability of a hospital to finance the treatments it provides out of the altruistic benefit it enjoys[22]. If that were not the case, the purchaser could rely on the goodwill of the hospital to ensure that treatments were carried out without the need for payment. We

[22] This is a similar formulation to Chone and Ma (2004) and Jack (2003). Jack (2003) studies the effect of private information regarding altruism on cost-reducing effort and quality choice and derives optimal cost sharing schemes. Chone and Ma (2004) show that private information regarding altruistic benefit can lead to pooling of patient types.

assume that $A(x, \phi)$ is such that the prices p_g and p_b remain positive in equilibrium.

Under full information, the effect of altruism is to improve the purchaser's welfare because by relying on the hospital's concern for its patients, the purchaser can reduce its payment to the hospital. The purchaser's objective function continues to be given by (1) whilst the individual rationality constraints IR_g and IR_b need to be modified to reflect the fact that the prices p_g and p_b do not have to cover the entire cost of treatment. Therefore, the constraints become

$$p_i - c(x_i, \phi_i) + A(x_i, \phi_i) \geq 0, \qquad (IR_i^A)$$

for $i = g, b$. Since these constraints are binding, purchasers can take full advantage of a hospital's altruism and reduce the payment for each type by the full amount of altruistic benefit $A(x_i, \phi_i)$. This implies that full-information treatment intensities will be higher under altruism.

Under asymmetric information, it is notable that the hospital's altruistic benefit is private information and therefore provides a source of informational rent. This means that the purchaser may not be able to take as full an advantage of the hospital's altruism as it can under full information. So, whilst the optimal treatment intensities can be higher, the incentive to misrepresent patient types for a given pair of treatment intensities may also be higher.

We introduce the incentive constraints, but as before, only the individual rationality constraint for the bad type IR_b^A and the incentive compatibility constraint for the good type IC_g^A are binding. We write the two binding constraints for treatment-based payments first,

$$n(\bar{x}(x_{gt}, x_{bt}))[p_{gt} - c(x_{gt}, \phi_g) + A(x_g, \phi_g)]$$
$$= n(\bar{x}(x_{bt}, x_{bt})) \left[p_{bt} - c(x_{bt}, \phi_g) + A(x_{bt}, \phi_g)\right] \qquad (IC_{gt}^A)$$

$$p_{bt} - c(x_{bt}, \phi_b) + A(x_{bt}, \phi_b) = 0, \qquad (IR_{bt}^A)$$

and then the same two binding constraints for outcomes-based payments,

$$n(\bar{x}(x_{go}, x_{bo}))[p_{go} - c(x_{go}, \phi_g) + A(x_{go}, \phi_g)]$$
$$= n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[p_{bo} - c(\hat{x}_{bo}, \phi_g) + A(\hat{x}_{bo}, \phi_g)\right], \qquad (IC_{go}^A)$$

$$p_{bo} - c(x_{bo}, \phi_b) + A(x_{bo}, \phi_b) = 0. \qquad (IR_{bo}^A)$$

From these, we can compute the rent under the two schemes to be

$$R_t^A(x_{bt}) = n(\bar{x}(x_{bt}, x_{bt})) \left[\left(c(x_{bt}, \phi_b) - c(x_{bt}, \phi_g)\right) + \left(A(x_{bt}, \phi_g) - A(x_{bt}, \phi_b)\right)\right], \quad (R_t^A)$$

and

$$R_o^A(x_{bo}) = n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[\left(c(x_{bo}, \phi_b) - c(\hat{x}_{bo}, \phi_g)\right) + \left(A(\hat{x}_{bo}, \phi_g) - A(x_{bo}, \phi_b)\right)\right]. \quad (R_o^A)$$

Under treatment-based payment, the rent from altruism, $\left(A(x_{bt}, \phi_g) - A(x_{bt}, \phi_b)\right)$, is positive indicating that the purchaser cannot take as full advantage of the hospital's altruism as it can under full information. However, under outcomes-based payment, the the rent expression, $\left(A(\hat{x}_{bo}, \phi_g) - A(x_{bo}, \phi_b)\right) = 0$ since $h(\hat{x}_{bo}, \phi_g) - h(x_{bo}, \phi_b)$. This implies that the analogous condition to (19), which establishes the superiority of outcomes-based schemes, is going to be less stringent in the presence of altruism.

We conclude that altruism will bias the purchaser towards outcomes-based payments. Since altruistic benefit is derived from the patient's health gain, outcomes monitoring constrains the hospital in benefiting from altruism.

This can also be seen formally by repeating the steps of the proof of proposition 3 and comparing the analogous expressions for the minimum payments p_{gt} and p_{go} that implement an identical pair of treatment intensities, x_g and x_b, under the two payment schemes. In these expressions, the only new parts due to altruism are the two rent terms $\left(A(x_b, \phi_g) - A(x_b, \phi_b)\right)$ for p_{gt} and $\left(A(\hat{x}_b, \phi_g) - A(x_b, \phi_b)\right)$ for p_{go}. It is immediately clear that the p_{go} will be reduced relative to p_{gt} when there is altruism.

8 Discussion

Third party purchasers of health service seeking to minimize cost must design purchasing arrangements to make use of all available information. In spite of being possibly ill-informed about their health status and insulated from the costs of services, patients have an important role to play in this process as a disciplining device that can be exploited by purchasers. In this paper we have demonstrated that monitoring schemes differ in their effectiveness in utilizing patients as a disciplining device and shown that the choice of the monitoring instrument may itself depend on the strength of patients' demand response to variations in quality of service. We have developed a framework for assessing two fundamental ways of calculating payments: one based on treatments (input measures), and one based on health outcomes (improvement in health status). We have shown that when demand is particularly responsive to quality of service, payment schemes based on outcome reduce the overall cost to the purchaser relative to payment schemes based on treatment. We show that this is because payment based on outcomes makes misrepresentation of patient-type more costly to a supplier in terms of a demand effect, since services will have to be tailored to patient-type such that the outcome is consistent with what is claimed. We have extended the analysis to consider the impact of service providers who have an intrinsic concern for patients and are, thus, altruistic. The obvious effect is that the purchaser can reduce payment to suppliers to take advantage of their altruism. Asymmetric information, however, prevents purchasers from exploiting altruism fully. Closer to the focus the paper, altruism biases purchasers towards outcome-based payments because misrepresentation hurts patients more under outcome-based payments and this impacts upon the supplier's utility.

The *outcomes movement* seeks to clarify what are appropriate treatments for different medical conditions by means of measuring the benefits from treatments and thus extends the information available to the purchasers of health care. In practice it seems likely that there will always be greater uncertainty associated with measuring outcome, where the object of analysis is an individual's health status, than with measuring treatment and thus that conditioning payment upon outcomes is inherently risky. Nevertheless, there is considerable ongoing investment in acquiring more and better outcomes information and our analysis indicates how this extra information might be used and, in particular, that there is a benefit to directly incorporating it into the payment of health care providers when, as is true for some conditions, patients actively choose whether to be treated according to the quality of service that they expect to receive. When hospitals are altruistic our analysis indicates that outcomes based payment is also more likely. Hence, according to both the type of health service being provided and the objectives of the provider, different purchasing arrangements are appropriate.

Insights from this analysis are applicable in other settings where third party purchasing occurs and clients do not clearly observe benefits received. Besides markets where insurance is important, third party purchasing also frequently occurs where public provision is undertaken and users of services are insulated, perhaps for reasons of equity, from the direct costs of the services they consume. In the UK many of the services that in the US are included under health care are termed social services and are provided on this basis. In both the UK (Legal Aid) and the US (public defenders) legal services are provided by the state also on this basis. Our findings suggest that input based payments will be more effective where demand results from an emergency, while outcome based payments may be more attractive in situations where consumers are more free to exercise choice.

Appendix

Proof of proposition 1

Proof. The result on rents follow from the binding constraints (7) and (8). Using the definitions of $S(.)$ from (2), rent from (9), and the fact that (7) and (8) hold as equalities, we can rewrite the purchaser's problem as the unconstrained problem,

$$\max_{x_{gt},x_{bt}} \left[n(\bar{x}(x_{gt},x_{bt}))S(x_{gt},x_{bt}) - \pi R_t(x_{bt}) \right], \tag{20}$$

or, by the definition of $W(.)$ from (3),

$$\max_{x_{gt},x_{bt}} \left[W(x_{gt},x_{bt}) - \pi R_t(x_{bt}) \right].$$

The first order conditions defining the optimal choices can be written,

$$W_g(\tilde{x}_{gt}, \tilde{x}_{bt}) = 0, \tag{21}$$

$$W_b(\tilde{x}_{gt}, \tilde{x}_{bt}) - \pi R_t'(\tilde{x}_{bt}) = 0 \tag{22}$$

where, using $\tilde{\cdot}$ to denote a function or derivative evaluated at the second best choices of treatment intensities $\tilde{x}_{gt}, \tilde{x}_{bt}$, we have

$$\begin{aligned} W_g(x_{gt}, x_{bt}) &= \tilde{n}'\pi\tilde{S} + \pi\tilde{n}\left(b_x(\tilde{x}_{gt}, \phi_g) - c_x(\tilde{x}_{gt}, \phi_g)\right), \\ W_b(x_{gt}, x_{bt}) &= \tilde{n}'(1-\pi)\tilde{S} + \tilde{n}(1-\pi)\left(b_x(\tilde{x}_{bt}, \phi_b) - c_x(\tilde{x}_{bt}, \phi_b)\right). \end{aligned}$$

Condition (21) implies that $b_x(\tilde{x}_{gt}, \phi_g) - c_x(\tilde{x}_{gt}, \phi_g) < 0$, and we know from the definition of $R_t(.)$ that $R_t'(\tilde{x}_{bt}) > 0$ since $n(\bar{x}(x_{bt}, x_{bt}))$ and $\left[c(\tilde{x}_{bt}, \phi_b) - c(\tilde{x}_{bt}, \phi_g)\right]$ are both increasing in x_{bt}.

I) $\tilde{x}_{bt} < \tilde{x}_{gt}$: The first order conditions (21) and (22) together imply that

$$\begin{aligned} \left(b_x(\tilde{x}_{gt}, \phi_g) - c_x(\tilde{x}_{gt}, \phi_g)\right) &= \left(b_x(\tilde{x}_{bt}, \phi_b) - c_x(\tilde{x}_{bt}, \phi_b)\right) - \frac{\pi}{\tilde{n}(1-\pi)} R_t'(\tilde{x}_{bt}) \\ &< \left(b_x(\tilde{x}_{bt}, \phi_b) - c_x(\tilde{x}_{bt}, \phi_b)\right). \end{aligned}$$

Therefore, $\tilde{x}_{gt} > \tilde{x}_{bt}$ since (a) $b_x(\tilde{x}_{gt}, \phi_g) - c_x(\tilde{x}_{gt}, \phi_g) < 0$, (b) $b_{xx}(.) - c_{xx}(.) < 0$, and (c) $\left(b_x(x, \phi_g) - c_x(x, \phi_g)\right) > \left(b_x(x, \phi_b) - c_x(x, \phi_b)\right)$ for all x.

II) $\tilde{x}_{bt} < x_b^*$: Consider a small change (dx_{gt}, dx_{bt}) and evaluate the objective function near the first best. Since $W_g(x_g^*, x_b^*) = W_b(x_g^*, x_b^*) = 0$, and $R_t'(\tilde{x}_{bt}) > 0$, the value of the objective function decreases with x_{bt} at (x_g^*, x_b^*). Therefore, $\tilde{x}_{bt} < x_b^*$.

III) We cannot determine whether $\tilde{x}_{gt} > or < x_g^*$ because the sign of the cross partial derivative W_{gb} is ambiguous in general. $\blacksquare$

Proof of proposition 2

Proof. The result on rents follows from the binding constraints (12) and (13). Using the definitions of $S(.)$ from (2) rent from (9), and the fact that (12) and (13) hold as equalities, we can rewrite the purchaser's problem as the unconstrained problem,

$$\max_{x_{go},x_{bo}} \left[n(\bar{x}(x_{go},x_{bo}))S(x_{go},x_{bo}) - \pi R_o(x_{bo})\right], \tag{23}$$

or, by the definition of $W(.)$ from (3),

$$\max_{x_{go},x_{bo}} \left[W(x_{go},x_{bo}) - \pi R_o(x_{bo})\right]$$

The first order conditions defining the optimal choices can be written,

$$W_g(\tilde{x}_{go},\tilde{x}_{bo}) = 0, \tag{24}$$

$$W_b(\tilde{x}_{go},\tilde{x}_{bo}) - \pi R_o'(\tilde{x}_{bo}) = 0 \tag{25}$$

The rest of the proof follows that of proposition 1. However,it is useful to note that there is an additional complication in the case of outcome-based payment, which is

$$\begin{aligned} \frac{\partial c(\hat{x}_{bo},\phi_g)}{\partial x_{bo}} &= c_x(\hat{x}_{bo},\phi_g)\frac{b_x(x_{bo},\phi_b)}{b_x(\hat{x}_{bo},\phi_g)} \\ &< c_x(\hat{x}_{bo},\phi_g). \end{aligned}$$

This implies that $R_o(.)$ is larger and increases faster than $R_t(.)$ for each x_b. $\blacksquare$

Multiple types

We sketch the arguments needed to show that the extension of our model to multiple types may be tedious, but is not problematic. Let us consider a model with, three types such that $\phi \in \{\phi_g, \phi_m, \phi_b\}$ and the respective probabilities for ϕ_i are π_i, $i = b, m, g$. The arguments generalize to the case of n types in a straightforward manner. The binding IR constraint is IR_b and the binding incentive constraints are for the hospital claiming a good type to be medium and a medium type to be bad. The binding IC_g constraints for the good type under the two schemes are respectively IC_{gt} and IC_{go}:

$$\begin{aligned} n(\overline{x}(x_{bt}, x_{mt}, x_{gt})) \left[p_{gt} - c(x_{gt}, \phi_g)\right] &= \\ &\quad n(\overline{x}(x_{bt}, x_{mt}, x_{mt})) \left[p_{mt} - c(x_{mt}, \phi_g)\right], \\ n(\overline{x}(x_{bo}, x_{mo}, x_{go})) \left[p_{go} - c(x_{go}, \phi_g)\right] &= \\ &\quad n(\overline{x}(x_{bo}, x_{mo}, \hat{x}_{mo})) \left[p_{mo} - c(\hat{x}_{mo}, \phi_g)\right], \end{aligned}$$

and the binding IC_m constraints for the medium type under the two schemes are respectively IC_{mt} and IC_{mo}:

$$\begin{aligned} n(\overline{x}(x_{bt}, x_{mt}, x_{gt})) \left[p_{mt} - c(x_{mt}, \phi_m)\right] &= \\ &\quad n(\overline{x}(x_{bt}, x_{bt}, x_{gt})) \left[p_{bt} - c(x_{bt}, \phi_m)\right], \\ n(\overline{x}(x_{bo}, x_{mo}, x_{go})) \left[p_{mo} - c(x_{mo}, \phi_m)\right] &= \\ &\quad n(\overline{x}(x_{bo}, \hat{x}_{bo}, x_{go})) \left[p_{bo} - c(\hat{x}_{bo}, \phi_m)\right], \end{aligned}$$

while the binding IR constraints for the bad type under the two schemes respectively are IR_{bt} and IR_{bo}:

$$\begin{aligned} p_{bt} - c(x_{bt}, \phi_b) &= 0, \\ p_{bo} - c(x_{bo}, \phi_b) &= 0. \end{aligned}$$

From these expressions we can compute rents for the medium type as follows:

$$R_{mt}(x_{bt}, x_{gt}) = n(\overline{x}(x_{bt}, x_{bt}, x_{gt}))\left[c(x_{bt}, \phi_b) - c(x_{bt}, \phi_m)\right], \tag{R_{mt}}$$

$$R_{mo}(x_{bo}, x_{go}) = n(\overline{x}(x_{bo}, \hat{x}_{bo}, x_{go}))\left[c(x_{bo}, \phi_b) - c(\hat{x}_{bo}, \phi_m)\right]. \tag{R_{mo}}$$

In the proof of proposition 3, we essentially compare the rents under the two schemes for identical treatment intensities. So, we could rewrite the above to be implementing the intensities $\{x_b, x_m, x_g\}$ in both cases and compare the rent. But it is obvious that, comparing the rents of the medium type is identical to comparing the rents for the good type in a two-type case, and the same results as in the paper will follow. The cost differential in the square brackets is larger for outcomes-based payments, while the demand n is smaller. So, we move on to compare the rents for the good type in the current three-type model, but before that we need the prices for medium type under treatment and outcomes-based schemes respectively. In obvious notation these are p_{mt} and p_{mo} :

$$\begin{aligned} p_{mt} &= \frac{R_{mt}(x_b, x_g)}{n(\overline{x}(x_b, x_m, x_g))} + c(x_m, \phi_m), \\ p_{mo} &= \frac{R_{mo}(x_b, x_g)}{n(\overline{x}(x_b, x_m, x_g))} + c(x_m, \phi_m), \end{aligned}$$

where we have used identical intensities for both schemes. Using these computed prices in the IC_g equations, we then get the rent expressions for the good type.

$$\begin{aligned}
R_{gt} &= n(\overline{x}(x_b, x_m, x_m)) \left[\frac{R_{mt}(x_b, x_g)}{n(\overline{x}(x_b, x_m, x_g))} + c(x_{mt}, \phi_m) - c(x_m, \phi_g) \right], \\
R_{go} &= n(\overline{x}(x_b, x_m, \hat{x}_m)) \left[\frac{R_{mo}(x_b, x_g)}{n(\overline{x}(x_b, x_m, x_g))} + c(x_m, \phi_m) - c(\hat{x}_m, \phi_g) \right].
\end{aligned}$$

Again, when $n' \approx 0$, (i.e., n is almost a constant) these imply that $R_{gt} < R_{go}$ since then $R_{mo} > R_{mt}$ and we know that $\left(c(x_m, \phi_m) - c(\hat{x}_m, \phi_g)\right) > \left(c(x_m, \phi_m) - c(x_m, \phi_g)\right)$ because $\hat{x}_m < x_m$. For a high enough n', the opposite is true. The cost expressions are not affected by $n'$and neither are the denominators $n(\overline{x}(x_b, x_m, x_g))$. But $[n(\overline{x}(x_b, x_m, x_m)) - n(\overline{x}(x_b, x_m, \hat{x}_m))]$ increases and we know from above that $R_{mt} > R_{mo}$ for n' high enough.

Hence the analysis in the paper generalizes in a natural way to three types and, by a recursion argument, to n types.

Acknowledgements

We would like to thank two anonymous referees of this journal, Ingela Alger, Doug Conrad, Jacques Lawarrée, Albert Ma, Jim Malcomson, Hugh Richardson, Eugene Silberberg and participants at the 4th Biennial Industrial Organization of Health Care Conference, Portsmouth N.H. for comments and suggestions. The first author gratefully acknowledges the support of the Economic and Social Research Council (ESRC). The work was partially funded by ESRC award number R000236723.

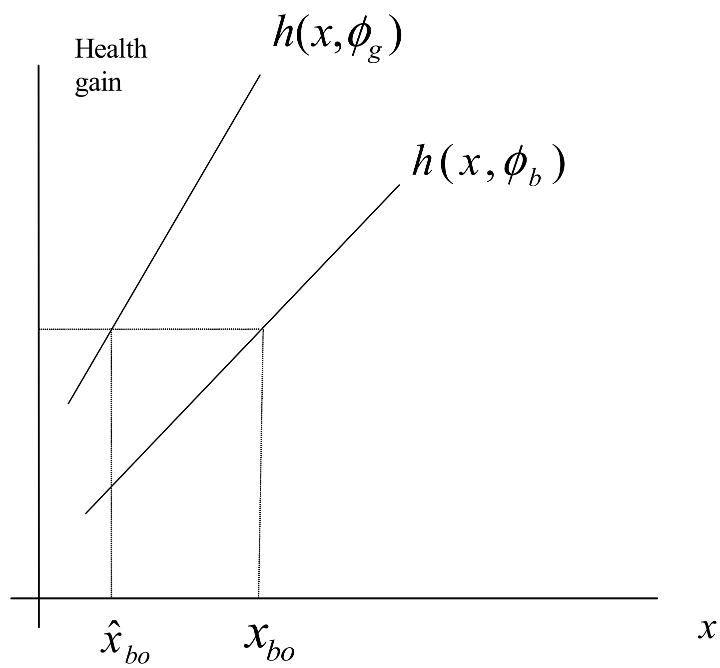

Figure 1: Treatment under outcome-based payment

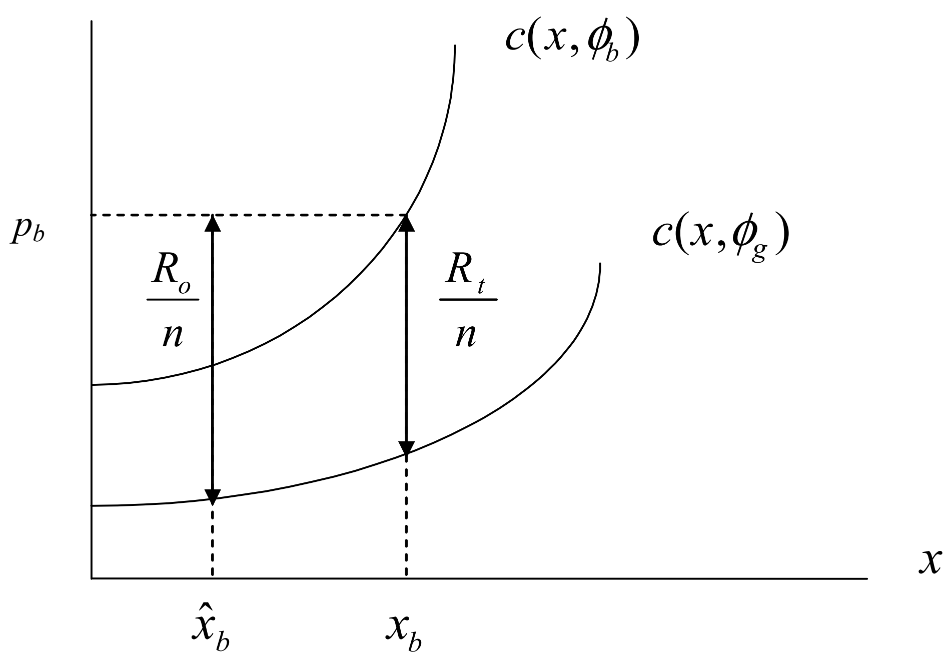

Figure 2: Rent per person under treatment and outcome-based payments.