OPEN ACCESS

UNIVERSITY OF DUNDEE

University of Dundee

Third party purchasing and incentives: the "Outcome Movement" and contracts for health services

Chalkley, Martin; Khalil, Fahad

Publication date:
2001

Link to publication in Discovery Research Portal

Dundee Discussion Papers in Economics

Third Party Purchasing and Incentives: The "Outcome Movement" and Contracts for Health Services

Martin Chalkley

and

Fahad Khalil

Third Party Purchasing and Incentives: The "Outcomes Movement" and Contracts for Health Services[1]

Martin Chalkley[2] Fahad Khalil[3]

31 January 2000

[2]Department of Economic Studies, University of Dundee, Dundee, DD1 4HN, U.K. *m.j.chalkley@dundee.ac.uk*

[3]Department of Economics, University of Washington, Seattle, WA 98115-3330. *khalil@u.washington.edu*

Abstract

The *Outcomes Movement*, also referred to as the "third revolution in medical care"
has focused attention and resources on assessing the effectiveness of medical treat-
ments. Outcomes research aims at determining the nature of effective treatment
regimes for medical conditions. Equipped with such information, health care pur-
chasers will be better able to regulate suppliers of health services. In this paper, we
consider whether payments conditioned on treatment delivered or on health outcome
obtained makes better use of the information provided by outcomes research. We
show that consumers can help third party purchasers allevaite incentive problems
due to asymmetric information. We find that, if patients in deciding whether to be
treated, are responsive to the treatments offered payment-by results may be preferred.
We further find that where health care suppliers operate not-for-profit, there is an
incentive for purchasers to specify payment-by results.

   JEL classification: I11

   Keywords: health, contracts, outcomes research, asymmetric information.

1 Introduction

Since the seminal contribution of Arrow (1963), it has been widely held that suppliers of health services are better informed regarding the effectiveness of medical treatments than either patients themselves or third party purchasers of health care and that this adversely affects the cost of health care provision. When the recipients of health services are insulated from the cost of services by insurance, they have an incentive to accept services provided only that those services yield some benefit. Since suppliers will deliver treatments provided that they are reimbursed for doing so, there is a tendency for costly treatments of limited value to be provided. This emphasis on the role of asymmetric information in increasing the costs of health care suggests that better informed purchasers will be better able to control the costs of health care by proscribing ineffective treatments. It is this premise that provides much of the impetus for what has been termed the *outcomes movement*[1].

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

---

[1]Definitions of outcomes research and a discussion of the role that it plays in health policy debates can be found in Agency for Health Care Policy and Research (1999) and UK Clearing House on Health Outcomes (1999).

[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

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 improvements in health status[5]. This activity has centered 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 analyzed. Proponents of this research see it as establishing the proper basis for measuring medical output[6] and, hence, claim that it will "bring order and predictability" (Ellwood, 1988) to health care systems. The rationale for such claims has 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.

In this paper we consider how purchasers of health services can make the *best* use of outcomes research in formulating their purchasing arrangements on the assumption that there is irreducible asymmetric information between purchasers and suppliers. Such asymmetry of information, which results in the need to pay rents to suppliers, seems inevitable. Indeed, disquiet with outcomes research in the medical profession

---

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.'

[5] See, for example, Tanenbaum (1993) who comments 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] This can be thought of as follows. Suppose that medical intervention has the effect of moving an individual from one health state to another. Since the outcome of treatment is never certain, the most that one can hope to learn is the probability of a particular treatment effecting a particular transition for a particular kind of ailment. Put this way the outcomes movement has as its objective the determination of the values in a set of Markov transition matrices describing movements between health states. The outcomes movement is both concerned with determining the appropriate labelling of the rows and columns of these transition matrices and with establishing the values of the probabilities themselves. If knowledge were ever to be complete, an individual patient could be described in terms of which transition matrices corresponded to their medical condition.

has centered on a concern that the knowledge of a particular patient acquired by a physician will always go beyond what can be described in terms of a patient's diagnosis and current health state[7].

If reliable data is available on health states, suppliers' payments can be based on such data as now happens in the case of some health services (Lu, 1999). In short, outcomes research facilitates 'payment-by-results' which might be used to elicit suppliers' hidden information. Alternatively, purchasers could use outcomes research to specify acceptable treatments and specify 'payment-by-treatment'. In this paper we compare the efficacy, in terms of reducing informational rents, of these alternatives. Henceforth, we refer to payment that is conditioned on what treatment achieves in terms of improving a patients' health as *outcome-based* and payment that is conditioned on the treatment given[8], as *treatment-based.* We show that the choice between basing payment on treatment or outcome depends on the cost of treatment, the attitude of suppliers to their patients and crucially on the extent to which patient-consumers respond to variations in the treatment offered when deciding where or whether to seek treatment.

The idea that patients who have a long term relationship with a physician or who receive information from friends regarding the treatment that is given by a particular supplier 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 a potentially important incentive instrument in mitigating the

---

[7]Tanenbaum (1993) argues that: "The knowledge that a procedure is more or less effective overall should not be confused with the knowledge of whether or how to use that procedure in caring effectively for the next patient. Whereas economists and clinical epidemiologists of the outcomes movement seek better clinical practice grounded in the objective certainties of statistical rigor, other students of medicine believe uncertainty and subjectivity are at the heart of the clinical encounter and insist that it will always be the case."

[8]Payments which are treatment specific are already used in some health care systems. McClellan (1997), for example, details examples of Diagnosis Related Groups (DRGs) in the US Medicare system which vary not according to the diagnosis but with the treatment given.

effects of *moral hazard* in health contracts; see for example Ma and McGuire (1997), Ma (1994), and Chalkley and Malcomson (1998a, 1998b, 2000).

We find that demand effects also help to align incentives by reducing provider rents due to asymmetric information. In addition, we show that outcome-based payment is more effective in utilizing patients' demand response. Therefore, when demand is particularly responsive, for example, for elective procedures, payment schemes based on outcome reduce the cost of health services 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 demand effect since treatments have to be tailored to patient-type such that the outcome is consistent with what is claimed.

These findings have a number of policy implications. First, they make clear the factors that influence the choice of health contracts in the light of greater medical knowledge and they provide insights into how health contracts can *best* be tailored to make use of that knowledge. Second they confirm the view that informed consumers have an important role to play in aiding third party purchasers control the costs of health services and therefore provide guidance as to where resources, targeted at improving consumers' information, might be best deployed.

The seminal work by Maskin and Riley (1985) establishes the theoretical framework concerning the choice of monitoring instruments in contracts with asymmetric information which is the foundation of this paper. Some recent contributions to that literature are Khalil and Lawarrée (1995), and Lewis and Sappington (1995). Our paper contributes to that literature as the results that we derive can be applied whenever a principal, who is a third party payer, contracts with a privately informed agent who faces a demand that is sensitive to the effort expended in production. The central insight we offer is that, with asymmetric information, demand responsiveness leads to an optimal contract in which payment is conditioned on the value of output

4

rather than effort input.

In the context of health care, 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*. Without the information that will (hopefully) be 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 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 $\phi_g > \phi_b$. A patient of type ϕ_g is a good prospect for treatment and will respond well to treatment and be cheap to treat. In the terminology of Whynes (1996), a treatment is both medically effective and cost effective for type ϕ_g, whereas the opposite is true for 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[10].

For each patient, the hospital determines an intensity of treatment[11], which we denote by x. A type-ϕ patient given a treatment of intensity x will have a gain in health status[12] of h, which is a random variable distributed with density $f(h \mid x, \phi)$, whose support is independent of x and ϕ, and has a mean increasing in x and ϕ. The purchaser attaches a monetary value $v(h)$ to the gain in health status h, and the purchaser's expected benefit is given by $b(x, \phi) \equiv \int v(h) f(h \mid x, \phi) dh$. We assume

---

[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]An alternative interpretation of asymmetric information in this context is that that there are not sufficient *risk adjusters* to fully reflect the differences between patients.

[11]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).

[12]We treat h as a scalar purely for convenience. It is straightforward to allow for a vector of characteristics representing an individual's health status.

that $b(x, \phi)$ is increasing and concave in x, increasing in ϕ and such that the marginal expected benefit of treatment is non-decreasing in ϕ. These assumptions ensure that the purchaser's expected benefit function is well behaved and meet the requirements 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 requirements of treatment being less costly for good types. 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). 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 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 is an increasing function of expected

---

[13]This, in common with much of the literature on health contracts, presumes that the health treatments being considered have the attributes of *search* goods.

intensity $\bar{x}$, and so denote total *expected* demand[14] 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 (possibly) 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.

Under what we call *treatment-based payment,* the transfer that the purchaser makes to the hospital depends on the distribution of types announced, which is checked *ex post* by verifying if the claimed treatments were provided. The purchaser pays p_{gt} for a patient claimed to be of type ϕ_g, i.e., when treatment x_{gt} is given, and it pays p_{bt} when a patient is claimed to be of type ϕ_b or for treatment of x_{bt}. We assume that the purchaser is able to impose a sufficient sanction such that the hospital has an incentive to provide the treatment it claims to have given to each patient.

Under what we call *outcome-based payment,* the transfer that the purchaser makes to the hospital again depends on the claimed distribution of types, but in this case claims are verified by establishing the health gain (i.e. the outcome) of a sample of patients treated perhaps using observable signals of health status. Hence, the purchaser pays p_{go} when a patient is claimed to be of type ϕ_g and it pays p_{bo} for a patient claimed to be of type ϕ_b. If, for example, the hospital claims that a proportion $1 - \pi^R$ of patients were type ϕ_b the purchaser can check whether the average health gain corresponds to $\overline{h}(\pi^R) \equiv \pi^R \int h f(h \mid x_{go}, \phi_g) dh + (1 - \pi^R) \int h f(h \mid x_{bo}, \phi_b) dh,$

---

[14]Random demand will make it infeasible to base penalties on the number of patients treated.

where the treatment intensities correspond with those that the hospital is required to provide for each type. We assume that there are sufficient numbers of patients treated and sampled such that the purchaser can verify the hospital's claim by evaluating the average health gain achieved. Again, we assume that the purchaser is able to impose sufficient sanction such that the verified health gains will indeed correspond to the claims made by the hospital.

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 given fixed cost of setting up a monitoring system necessary to verify a hospital's claims for payment, but that instituting a system to verify both treatment and outcomes is prohibitively expensive. In practice the cost of verifying outcomes or treatments could be different 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 incentive effects of one instrument at a time.

The timing of events that we assume is as follows: 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 and receive treatment, and the purchaser assesses

either the intensity of treatment that a patient undergoes or the outcome of treatment. The hospital is then paid according to its contract with the purchaser.

3   Full Information

If there is full information on patient types, the purchaser can condition payment directly on a patient's type. The purchaser's objective function[15] 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'[16]. We therefore consider the individual rationality constraints:

$$
p_g - c(x_g, \phi_g) \ \geq \ 0, \tag{IR_g}
$$

$$
p_b - c(x_b, \phi_b) \ \geq \ 0. \tag{IR_b}
$$

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[17]. To simplify notation, we define expected

---

[15] 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.

[16] 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.

[17] 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 $c(.)$

surplus per patient as

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

and write the purchaser's objective function as

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

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:

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

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

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

---

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 $c(.)$.

$W_{gb}(x_g^*, x_b^*)$ 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 net marginal benefit is negative 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 not 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 $\{p_{gt}, p_{bt}, x_{gt}, x_{bt}\}$. The optimal contract is the solution to the purchaser's problem P_t written below.

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

s. t.

$$n(\bar{x}(x_{gt}, x_{bt})) \left[p_{gt} - c(x_{gt}, \phi_g) \right] \geq n(\bar{x}(x_{bt}, x_{bt})) \left[p_{bt} - c(x_{bt}, \phi_g) \right], \tag{6}$$

$$p_{bt} - c(x_{bt}, \phi_b) \geq 0. \tag{7}$$

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 (6) and (7) will hold as equalities since the payment to the

hospital can otherwise be lowered to the purchaser's benefit. The left hand side of (6) measures the total rent that the hospital will earn under a treatment-based payment. Substituting from (7) into the right hand side of (6) and imposing equality, we can obtain an expression for this rent[18] as:

$$
\begin{aligned}
R_t(x_{bt}) &= n(\bar{x}(x_{bt}, x_{bt})) \left[c(x_{bt}, \phi_b) - c(x_{bt}, \phi_g) \right], \quad (8) \\
&> 0.
\end{aligned}
$$

In the solution[19] 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 (8). 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

---

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

[19]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.

$\left[c(x_{bt}, \phi_b) - c(x_{bt}, \phi_g) \right]$ 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. But, this can imply that x_{gt} will be greater than in the first best even though $R_t(.)$ is independent of x_{gt}. 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 to the good types.

Since $\tilde{x}_{gt} > \tilde{x}_{bt}$, demand responsiveness acts as a disciplining device because, as can be seen from expression (8), as the hospital adjusts treatment so as to elicit higher payment it must do so in a way that will cause consumers to abstain from being treated, which lowers total rent.

5 Outcome-based payment

Outcomes 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. If the hospital wants to misrepresent a good type as bad, it must provide a good type with an intensity of treatment such that the expected health gain is the same as if the patient was a bad type. Specifically, this intensity of treatment is

$$\hat{x}_{bo} = \hat{x}(x_{bo}, \phi_g, \phi_b), \tag{9}$$

defined by

$$\int h f(h \mid x_{bo}, \phi_b) dh \stackrel{\text{def}}{=} \int h f(h \mid \hat{x}_{bo}, \phi_g) dh. \tag{10}$$

The purchaser chooses the contract $\{p_{go}, p_{bo}, x_{go}, x_{bo}\}$ to solve the problem P_o given below.

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

s. t.

$$n(\bar{x}(x_{go}, x_{bo})) \left[p_{go} - c(x_{go}, \phi_g) \right] \geq n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[p_{bo} - c(\hat{x}_{bo}, \phi_g) \right], \quad (11)$$

$$p_{bo} - c(x_{bo}, \phi_b) \geq 0, \quad (12)$$

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

$$R_o(x_{bo}) = n(\bar{x}(\hat{x}_{bo}, x_{bo})) \left[c(x_{bo}, \phi_b) - c(\hat{x}_{bo}, \phi_g) \right], \quad (13)$$
$$> 0.$$

.

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   Comparing treatment and outcome-based payments

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_{bt} *in (9).*

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 (7) and (6), we have

$$p_{bt} = c(x_b, \phi_b), \tag{14}$$

$$p_{gt} = c(x_g, \phi_g) + \frac{n(\bar{x}(x_b, x_b))}{n(\bar{x}(x_g, x_b))} \left[c(x_b, \phi_b) - c(x_b, \phi_g) \right]. \tag{15}$$

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

$$p_{bo} = c(x_b, \phi_b), \tag{16}$$

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

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

$$n(\bar{x}(x_b, x_b)) \left[c(x_b, \phi_b) - c(x_b, \phi_g) \right] > n(\bar{x}(\hat{x}_b, x_b)) \left[c(x_b, \phi_b) - c(\hat{x}_b, \phi_g) \right]. \tag{18}$$

∎

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. The first effect is illustrated in Figure 1. In the Figure, the rent per good type patient under treatment-based payment (denoted R_t/n) is less than the rent per patient under outcome-based payment (denoted R_o/n). The difference will be greater the larger is the discrepancy be-

tween x_b and $\hat{x}_b$ and the greater is the discrepancy in the marginal cost of treatment intensity. 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. It follows from (18) 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 (18) 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 which is at odds with the empirical evidence, such as that presented by Dranove and White (1994), which indicates that not all hospitals are profit maximizers. 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[20]

---

[20]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. If that were not the case,

component to its objective function of $A(x, \phi) = \int a(h) f(h \mid x, \phi) dh$, where $a(h)$ is the hospital's valuation of h. We assume that $a(h) < v(h)$, so that the hospital does not fully reflect the purchaser's concerns for patient welfare and that $A(.)$ has the same curvature properties as $b(.)$.

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 for any chosen treatment intensities. The purchaser's objective function continues to be given by (1) whilst the constraints IR_g and IR_b need to be modified to reflect the fact that the prices p_g, p_b do not have to cover the entire cost of treatment. Therefore, the constraints become

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

for $i = g, b$. Under full information the payment for each type can be reduced by the full amount of altruistic benefit, which implies that the full-information level of treatment intensities will be higher than in the absence of altruism.

Under asymmetric information, the effect of altruism is more complex. The altruistic benefit that the hospital enjoys is private information and therefore provides a source of informational rent. When a good type is claimed to be bad and given a treatment of x_b, the benefit from altruism is $A(x_b, \phi_g)$, which is strictly greater than $A(x_b, \phi_b)$. Since the payment p_b must account for altruistic benefit of the amount $A(x_b, \phi_b)$, the hospital can retain some of the benefit when treating good types, specifically $(A(x_b, \phi_g) - A(x_b, \phi_b))$ per patient. Hence, the purchaser cannot take full advantage of the a hospital's altruism when there is asymmetric information about patient type.

---

the purchaser could rely on the goodwill of the hospital to ensure that treatments were carried out without the need for payment.

More pertinent to our analysis, altruism will bias the purchaser towards outcome-based payment. The reason is as follows. Under outcome-based payment, in order to misreport a good type as bad, the good type must be treated with lower intensity $(\widehat{x}_{bo})$ than a bad type (x_{bo}) and hence the amount of altruistic benefit that a hospital can retain is smaller. This implies that the price of treatment can be lowered more under outcome-based than under treatment-based payment.

8 Discussion

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. In this context, we have considered the relative merits of treatment-based and outcome-based payments when there is irreducible asymmetric information regarding the precise medical condition of a patient. With such asymmetric information the cost to the purchaser of ensuring the provision of treatment cannot be driven down to the cost of providing that treatment because there is informational rent. However, we have shown that outcome and treatment-based payment schemes differ in their ability to contain this additional cost of health care.

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 consumers have regarding the nature of medical interventions specifically when consumers use this information to decide whether to be treated. With consumers whose demand is responsive to treatment, a supplier of services that adjusts treatment so as to elicit higher payment must do so in a way that will cause consumers 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 value the health gain enjoyed by the patients that they treat. Then, a purchaser can again use outcome-based payment to impose a greater penalty on the supplier who distorts the treatment they offer in order to elicit higher payment. Clearly, other factors influence the choice of outcome versus treatment-based payment and the analysis shows that costs of treatment have an important role to play. The relative advantage of outcome-based payment increases the smaller is the marginal cost of treatment for those patients who are easy to treat and benefit most from treatment.

Our analysis has a number of implications for health care policy. The emergence of new knowledge regarding the effectiveness of medical treatments is trumpeted as paving the way for improvements in health care systems. By considering how outcomes research might be used in structuring health contracts, our research begins the process of examining the way in which improvements might be implemented. In our analysis outcomes research will reduce the cost of health services as borne by third party purchasers. But the analysis suggests that if the reduction in cost is to be as great as possible, purchasers will need to consider carefully how they base their payments – in particular whether payment is conditioned on the treatment given or the outcome achieved. The research further indicates those factors that will be crucial in making this choice and that the choice will vary according to the kinds of health services that are being purchased, the objectives of suppliers and the extent to which consumers are well-informed about treatments on offer. Consistent with other studies of health contracts that have focussed on moral hazard issues, we find that the role of consumers who respond by demanding treatment according to their perceptions of

the treatment that is on offer, is crucial. Our research, therefore, complements those studies of health contracts under moral hazard by suggesting an equivalent role for informed consumers in mitigating the effects of asymmetric information and suggests that policy directed at providing consumers with better information may be valuable in containing the costs of health care.

In practice, treatment-based and outcome-based payments schemes will require very different methods of implementation. For a treatment-based payment, the purchaser will need to monitor treatments actually given so as to ensure that these meet the requirements of payment. Whilst this may be difficult and hence costly, the problems encountered would appear less than when it is necessary to monitor health outcomes. Outcomes research is still at a formative stage and there is as yet no clear consensus on what the relevant dimensions of outcome measurement are – at least not for the great majority of health interventions. We have not in this paper considered these practical issues, which will require further research before outcome or treatment-based payments schemes can be effectively implemented. This observation does not however preclude the value of exploring under precisely which conditions outcome-based payment *might* be beneficial and it is in that spirit that we present our results in the context of contracts for health services.

The central insight of our analysis also carries across to different contractual settings. Away from the application to health markets considered here, the approach that we adopt can be applied whenever a principal who is a third party payer contracts with an agent who faces a demand for their goods or services that is sensitive to the effort expended in production. The central insight offered here is that demand responsiveness on the part of consumers leads a principal towards monitoring the value of output rather than the monitoring of effort inputs.

23

9 Appendix

9.1 Proof of proposition 1

Proof. The result on rents follow from the binding constraints (6) and (7). Using the definitions of $S(.)$ from (2), rent from (8), and the fact that (6) and (7) hold as equalities, we can rewrite the purchaser's problem as the unconstrained problem,

$$\max_{x_{gt}, x_{bt}} \left[n(\bar{x}(x_{gt}, x_{bt})) S(x_{gt}, x_{bt}) - \pi R_t(x_{bt}) \right], \tag{19}$$

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,

$$W_g(\tilde{x}_{gt}, \tilde{x}_{bt}) = 0, \tag{20}$$

$$W_b(\tilde{x}_{gt}, \tilde{x}_{bt}) - \pi R_t'(\tilde{x}_{bt}) = 0 \tag{21}$$

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

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

Condition (20) 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 (20) and (21) together imply that

$$\left(b_x(\tilde{x}_{gt}, \phi_g) - c_x(\tilde{x}_{gt}, \phi_g)\right) = (b_x(\tilde{x}_{bt}, \phi_b) - c_x(\tilde{x}_{bt}, \phi_b)) - \frac{\pi}{\tilde{n}(1-\pi)} R'_t(\tilde{x}_{bt})$$

$$< (b_x(\tilde{x}_{bt}, \phi_b) - c_x(\tilde{x}_{bt}, \phi_b)).$$

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) $\left(b_x(x, \phi_g) - c_x(x, \phi_g)\right) > (b_x(x, \phi_b) - c_x(x, \phi_b))$ 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. ∎

9.2   Proof of proposition 2

Proof. The result on rents follows from the binding constraints (11) and (12). Using the definitions of $S(.)$ from (2) rent from (8), and the fact that (11) and (12) hold as equalities, we can rewrite the purchaser's problem as the unconstrained problem,

$$\max_{x_{go}, x_{bo}} \left[n(\bar{x}(x_{go}, x_{bo}))S(x_{go}, x_{bo}) - \pi R_o(x_{bo})\right], \tag{22}$$

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,

$$W_g(\tilde{x}_{go}, \tilde{x}_{bo}) = 0, \tag{23}$$

$$W_b(\tilde{x}_{go}, \tilde{x}_{bo}) - \pi R_o'(\tilde{x}_{bo}) = 0 \tag{24}$$

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. ∎

References

Agency for Health Care Policy and Research (1999). Outcomes research: Fact sheet, Internet http://www.ahcpr.gov/clinic/outfact.htm.

Allen, R. and Gertler, P. (1991). Regulation and the provision of quality to heterogenous consumers: The case of prospective pricing of medical services, *Journal of Regulatory Economics* **3**: 361–375.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care, *American Economic Review* **53**(5): 941–973.

Chalkley, M. and Malcomson, J. M. (1998a). Contracting for health services when patient demand does not reflect quality, *Journal of Health Economics* **17**(1): 1–19.

Chalkley, M. and Malcomson, J. M. (1998b). Contracting for health services with unmonitored quality, *Economic Journal* **108**(449): 1093–1110.

Chalkley, M. and Malcomson, J. M. (2000). Government purchasing of health services, *in* A. Culyer and J. P. Newhouse (eds), *Handbook of Health Economics*, North-Holland, Amsterdam.

Dranove, D. (1987). Rate-setting by diagnosis related groups and hospital specialization, *RAND Journal of Economics* **18**(3): 417–427.

Dranove, D. and White, W. D. (1994). Recent theory and evidence on competition in hospital markets, *Journal of Economics and Management Strategy* **3**(1): 169–209.

Ellis, R. P. and McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply, *Journal of Health Economics* **5**: 129–151.

Ellwood, P. M. (1988). Outcomes management: A technology of patient experience, *New England Journal of Medicine* **318**: 1549–1556.

Epstein, A. M. (1990). The outcomes movement - will it get us where we want to go?, *New England Journal of Medicine* **323**: 266–270.

Khalil, F. and Lawarrée, J. (1995). Input versus output monitoring, *Journal of Economic Theory* **66**: 139–157.

Laffont, J.-J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge, MA.

Lewis, T. R. and Sappington, D. E. M. (1995). Using markets to allocate pollution permits and other scarce resource rights under limited information, *Journal of Public Economics* **57**(3): 431–455.

Lewis, T. R. and Sappington, D. E. M. (1999). Using subjective risk adjusting to prevent patient dumping in the health care industry, *Journal of Economics and Management Strategy* **8**(3): 351–382.

Lu, M. (1999). Separating the true effect from gaming in incentive-based contracts in health care, *Journal of Economics and Management Strategy* **8**(3): 383–431.

Ma, C.-t. A. (1994). Health care payment systems: Cost and quality incentives, *Journal of Economics and Management Strategy* **3**(1): 93–112.

Ma, C.-t. A. and McGuire, T. G. (1997). Optimal health insurance and provider payment, *American Economic Review* **87**(4): 685–704.

Maskin, E. and Riley, J. (1985). Input versus output incentive schemes, *Journal of Public Economics* **28**: 1–23.

McClellan, M. (1997). Hospital reimbursement incentives: An empirical analysis, *Journal of Economics and Management Strategy* **6**(1): 91–128.

Naylor, C. D. (1995). Grey zones of clinical practice : Some limits of evidence based medicine, *Lancet* **345**: 840–842.

Newhouse, J. P. (1970). Toward a theory of nonprofit institutions: An economic model of a hospital, *American Economic Review* **60**(1): 64–74.

Relman, A. S. (1988). Assessment and accountability: The third revolution in medical care, *New England Journal of Medicine* **319**: 1220–1222.

Tanenbaum, S. J. (1993). What physicians know, *New England Journal of Medicine* **329**: 1268–1271.

UK Clearing House on Health Outcomes (1999). About health outcomes, Internet http://www.leeds.ac.uk/nuffield/infoservices/UKCH/about.html.

Whynes, D. K. (1996). Towards an evidence-based National Health Service, *Economic Journal* **106**(439): 1702–1712.