



DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

**SUPPLIER DISCRETION OVER PROVISION: THEORY AND AN
APPLICATION TO MEDICAL CARE**

James M. Malcomson

Number 221

January 2005

Manor Road Building, Oxford OX1 3UQ

Supplier Discretion over Provision: Theory and an Application to Medical Care

James M. Malcomson*
University of Oxford, U.K.

First version: May 2003
This version: January 10, 2005

*I would like to thank Philippe Choné, Jacques Crémer, Alastair Gray, Meg Meyer, Nolan Miller, participants in the Economic Theory Workshop at Oxford and in the 6th Conference on the Industrial Organization of Health Care held at Hyannis in April 2004, and two anonymous referees for very helpful comments. The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The research was funded by ESRC award number R000236723.

Abstract

Suppliers who are better informed than purchasers, such as physicians treating insured patients, often have discretion over what to provide. This paper shows how, when the purchaser observes what is supplied but neither recipient type nor the actual cost incurred, optimal provision differs from what would be efficient if the purchaser had full information, whether or not the supplier can extract informational rent. The analysis is applied to, among other things, data on tests for coronary artery disease and to Medicare *diagnosis-related groups* defined by the treatment given, not just the diagnosis, illustrating the biases in provision that result.

Keywords: Supplier discretion, Procurement, Public provision, Diagnosis-related groups, Medicare, Prospective payment, Cost-effectiveness

JEL classification: D82, H42, I11, I18

1 Introduction

Medical specialists decide which treatment to provide to patients — whether, for example, to treat coronary artery disease surgically with a bypass graft or non-surgically. Auto repairers have discretion about repairing or replacing parts of fully-insured autos damaged in an accident. Firms follow the advice of specialist suppliers about what computers to provide for employees. Providers of social care decide what services to supply to welfare recipients. Schools decide what special needs provisions to make for individual pupils. In all these cases, the choice of precisely what is to be supplied is left to the discretion of the supplier, with payment based on what the supplier provides. This paper is concerned with payment arrangements in such settings.

The settings analysed here differ from the procurement models in Laffont and Tirole (1993) in two fundamental respects. The first is the supplier's discretion over precisely what to supply. In some cases, such as emergency medical care, that discretion arises from the need for a speedy decision. But even where it is feasible to give the discretion to an agent employed by the purchaser (for example, a health service gatekeeper), employing an agent costs money and has its own incentive problems, see Malcomson (2004). Whatever the reason, in practice discretion is often left to the supplier. The second important difference is that payment to the supplier is not conditioned on the actual cost incurred in supplying each individual recipient. An important practical reason is the difficulty of monitoring the cost actually incurred. With health services, this was a major motivation for the shift from *fee-for-service* payment, based essentially on the cost of the services actually supplied to a patient, to *prospective* payment, based on *diagnosis-related groups* (DRGs) defined by the patient's basic diagnosis and also frequently, as McClellan (1997) emphasises, by the specific procedure used to treat the patient.

When the supplier has discretion, the purchaser paying for provision needs to consider how payment arrangements affect not only the cost, but also the appropriateness, of provision. To investigate the implications, this paper uses a framework with two varieties of provision (treatment, computer system, social care service, etc.) that might be appropriate for a recipient. Recipients differ in type, which is observed by the supplier before deciding which (if either) of the two varieties to supply but not observed by the purchaser. For some recipient types, one variety of provision is efficient (in the sense of maximising the net benefit as assessed by the purchaser), for some the other, and for some neither. The purchaser can verify which variety is provided for each recipient but not the cost incurred in providing it. Recipients make no payments to the supplier. Those made by the purchaser can be conditioned only on the variety supplied.

Two special cases of this framework are the following. First, the purchaser can be the same as the recipient provided type is not known to the recipient. In many of the applications, however, these two are distinct so the distinction is retained here. Second, type can be a cost characteristic of a monopoly supplier unknown to the purchaser. Then the model is similar in spirit to Baron and Myerson (1982) with output restricted to three possible levels, though

the setting is in other respects richer. Armstrong and Sappington (2005) survey the literature developed from that paper.

Critical for determining what provision is optimal in this setting are two properties of the costs of provision. The first property is whether there are recipient types for whom neither variety is efficient who are cheaper to supply than those for whom provision of some variety is efficient. Because efficient provision depends on benefits as well as costs, that can certainly be the case. If there are, these types will always receive provision if other types do, so there is necessarily some distortion away from efficient provision. If there are not, the purchaser can ensure that only those for whom some variety is efficient receive provision by setting the payments for providing each variety appropriately. The second property concerns how the difference in the cost of supplying the two varieties changes with recipient type. If it changes in one direction, an appropriate difference in payment between the two varieties can ensure that those for whom some provision is efficient receive the variety that is efficient for them. If it changes in the other direction, inducing the supplier to respond in that way is simply not feasible. If the supplier can be prevented from receiving a rent from its information about recipient types, the difference between optimal and efficient provision is determined entirely by these constraints on feasibility. If, however, the supplier cannot be prevented from receiving such a rent, optimal provision is adjusted in order to reduce that rent. This may result in some types receiving the variety that is efficient for them when that would not otherwise be the case.

I am aware of only a small literature on settings similar to that analysed here, all of it concerned specifically with DRGs for different medical treatments for the same basic diagnosis. McClellan (1996a) and McClellan (1996b) both consider two possible treatments given the basic diagnosis. Neither is concerned with the optimal choice of DRG prices. Instead, they are concerned with the implications of DRG prices that follow a specified (and not, in general, optimal) rule on hospital investment in capacity to carry out the treatments, on physicians' allocation of patients to the treatments, and on patients' choices of which hospital to attend. The results here, in contrast, are concerned with prices that are optimal for the purchaser given the rule for allocating provision to recipients that is optimal for suppliers. Miller (2004) also analyses a setting with two potential treatments but his concern is with the structure of medical insurance, specifically whether it is more efficient to have insurance provided by the health maintenance organisation supplying treatment or by a third-party insurer. His conclusion is that it is more efficient to have a third-party insurer, the case analysed here of a purchaser separate from the supplier. Siciliani (2003) too analyses a setting with two potential treatments but under very specific assumptions about costs of treatment (increasing linear cost functions), benefits of treatment (the same for both treatments and all patient types), and the distribution of patient types (uniform). Those assumptions correspond to a special case of one of the four cases considered here. His primary concern is with how optimal DRG prices vary across hospitals with different distributions of patient types.

The paper is organised as follows. The next section sets out the model used for the analysis. Sections 3 and 4 derive results on optimal prices for the two varieties of provisions for

different cost configurations. Section 5 applies the results to medical care. The basic issue there is how the cost and benefit characteristics relevant for the economic analysis relate to the characteristics of medical treatments. Section 6 discusses other applications.

2 The model

Potential recipients of provision differ in type, indexed by the parameter $s \in [\underline{s}, \bar{s}]$, that affects what provision (if any) is appropriate. The distribution of types in the population served by the supplier is $F(s)$, with corresponding density function $f(s) > 0$ for all $s \in (\underline{s}, \bar{s})$. Two varieties of provision are available, variety l appropriate for low s recipients and variety h appropriate for high s recipients. The cost of providing variety i to a recipient of type s is $c_i(s) > 0$ for $i \in \{l, h\}$, which is differentiable. (Appendix A extends the model to the case in which costs are affected by the supplier's choice of effort.) The cost functions and the distribution of types are known to both purchaser and supplier. In addition, before arranging for supply, the supplier observes a recipient's type but this remains unknown to the purchaser, as does the cost actually incurred in supplying the recipient. The purchaser can, however, monitor costlessly which variety of provision has been supplied. There may in addition be some types $s \notin [\underline{s}, \bar{s}]$ for which the appropriate provision is so obviously apparent that no supplier would consider giving the less appropriate one.

The monetary value of the benefit ascribed by the purchaser to a type s recipient receiving provision of variety i is denoted $b_i(s)$, which is positive and differentiable. For a public sector purchaser of health or social care, this may be simply the benefit to the recipient. For a firm buying computers, it is the increased profits that result. The formulation is consistent with the purchaser being the recipient if the benefit is independent of s so that s is not revealed to the purchaser once the benefit is known. A purchaser paying from public funds attaches a premium $\alpha > 0$ to payments, so the social cost of provision is $(1 + \alpha)c_i(s)$. For a private sector purchaser, $\alpha = 0$. The net benefit of provision is thus $b_i(s) - (1 + \alpha)c_i(s)$. To capture the property that variety l is appropriate for low s and variety h for high s , types s are ordered such that there exist $s' \in (\underline{s}, \bar{s})$ and $s'' \in (s', \bar{s})$ satisfying

$$b_l(s) - (1 + \alpha)c_l(s) > \max[0, b_h(s) - (1 + \alpha)c_h(s)], \text{ for } s \in [\underline{s}, s'), \quad (1)$$

$$b_h(s) - (1 + \alpha)c_h(s) > \max[0, b_l(s) - (1 + \alpha)c_l(s)], \text{ for } s \in (s', s''), \quad (2)$$

$$b_l(s) - (1 + \alpha)c_l(s) < b_h(s) - (1 + \alpha)c_h(s) < 0, \text{ for } s \in (s'', \bar{s}]. \quad (3)$$

The term *efficient provision* is used here for the provision that maximises the net benefit to the purchaser. Conditions (1)–(3) ensure that variety l is efficient for recipients of type $s \in [\underline{s}, s')$, variety h is efficient for recipients of type $s \in (s', s'')$, and neither variety is efficient for recipients of type $s \in (s'', \bar{s}]$. Either l or h is efficient for $s = s'$, either h or no provision for $s = s''$. If it is efficient for all types $s \in [\underline{s}, \bar{s}]$ to receive one or other provision, then $s'' = \bar{s}$, in which case it is assumed that $b_h(s'') - (1 + \alpha)c_h(s'') = 0$.

As typically assumed in procurement models, the purchaser makes a “take it or leave it” contract offer to the supplier. In the present case, a contract $P = \{p_0, p_l, p_h\}$ consists of prices p_i that the purchaser pays the supplier for providing a recipient with variety $i \in \{l, h\}$ and a payment p_0 for each potential recipient, whether or not supplied. (By the taxation principle, see Guesnerie (1995, Chapter 1), just one price per provision is a characteristic of an optimal direct mechanism.) The payment p_0 may, in principle, be negative but then the supplier has no interest in notifying the purchaser of a potential recipient who receives no provision. A purchaser who cannot monitor these is effectively constrained to set $p_0 \geq 0$. Both possibilities are analysed in what follows. The supplier’s payoff from providing variety i to a recipient of type s is

$$u_i(s) = p_i + p_0 - c_i(s). \quad (4)$$

The supplier’s reservation payoff is normalised to zero.¹

Let $S_i(P)$ denote the subset of types provided with variety $i \in \{l, h\}$ when the contract is P . The supplier’s participation constraint can then be written

$$p_0 + \int_{S_l(P)} [p_l - c_l(s)] f(s) ds + \int_{S_h(P)} [p_h - c_h(s)] f(s) ds \geq 0. \quad (5)$$

Since p_0 is paid whether or not provision is made, it affects provision only through this participation constraint. Moreover, since the supplier provides variety $i \in \{l, h\}$ to type s only if $p_i - c_i(s) \geq 0$, the constraint $p_0 \geq 0$, when imposed, is always tighter than (5).

It is convenient to nest the objective functions of profit-maximising and social-welfare-maximising purchasers within a single formulation. For a profit-maximising purchaser, the gain from having variety $i \in \{l, h\}$ provided to a recipient of type s is the benefit less the price paid, $b_i(s) - (p_i + p_0)$. The gain from no provision being supplied to a potential recipient is $-p_0$. The purchaser’s payoff from provision to a recipient of unknown type is, therefore,

$$W^\pi = \int_{S_l(P)} [b_l(s) - p_l] f(s) ds + \int_{S_h(P)} [b_h(s) - p_h] f(s) ds - p_0. \quad (6)$$

It is clearly optimal to set p_0 at the lowest level consistent with the participation constraint (5) and the constraint $p_0 \geq 0$ when the latter is required. Since the latter is a tighter constraint, it holds with equality when imposed. Otherwise, the participation constraint (5) holds with equality and can be used to substitute for p_0 in (6). To cover both cases, let $\delta = 1$ when p_0 is constrained to be non-negative and $\delta = 0$ when p_0 is not so constrained so that expected profit

¹The results that follow are easily extended to a supplier whose payoff depends on the benefit of provision to the recipient, as sometimes argued for physicians and for non-profit suppliers, see Newhouse (1970) and McGuire (2000). If the supplier’s payoff is $u_i(s) = p_i - v_i[\bar{c}_i(s), b_i(s)]$, where $\bar{c}_i(s)$ is the monetary cost of providing variety i to type s and $b_i(s)$ the benefit, the results that follow still apply when $c_i(s)$ is defined as $v_i[\bar{c}_i(s), b_i(s)]$.

W^π in (6) can be written

$$\begin{aligned}
W^\pi &= \int_{S_l(P)} [b_l(s) - p_l] f(s) ds + \int_{S_h(P)} [b_h(s) - p_h] f(s) ds \\
&\quad + (1 - \delta) \left[\int_{S_l(P)} [p_l - c_l(s)] f(s) ds + \int_{S_h(P)} [p_h - c_h(s)] f(s) ds \right] \\
&= \int_{S_l(P)} [b_l(s) - (1 - \delta) c_l(s) - \delta p_l] f(s) ds \\
&\quad + \int_{S_h(P)} [b_h(s) - (1 - \delta) c_h(s) - \delta p_h] f(s) ds.
\end{aligned} \tag{7}$$

For a social-welfare-maximising purchaser, the gain from having variety $i \in \{l, h\}$ provided to a recipient of type s , denoted $w_i(s)$, consists of the benefit $b_i(s)$ to the recipient, plus the utility $u_i(s)$ to the supplier, less the social cost of paying for provision from public funds. Thus,

$$\begin{aligned}
w_i(s) &= b_i(s) + u_i(s) - (1 + \alpha)(p_i + p_0) \\
&= b_i(s) - c_i(s) - \alpha(p_i + p_0),
\end{aligned} \tag{8}$$

the second equality following from the definition of $u_i(s)$ in (4). For a potential recipient for whom no provision is made social welfare is $-\alpha p_0$. Thus the expected social welfare from providing for a recipient of unknown type is

$$W^{sw} = \int_{S_l(P)} [b_l(s) - c_l(s) - \alpha p_l] f(s) ds + \int_{S_h(P)} [b_h(s) - c_h(s) - \alpha p_h] f(s) ds - \alpha p_0. \tag{9}$$

Since $\alpha > 0$ for such a purchaser, it is again optimal to set p_0 at the lowest level consistent with the participation constraint (5) and the constraint $p_0 \geq 0$ when the latter is required. Since, as with a profit-maximising purchaser, the latter is a tighter constraint, it holds with equality when imposed. Otherwise, the participation constraint (5) holds with equality and can be used to substitute for p_0 in (9). In this case, let $\delta = \alpha$ when p_0 is constrained to be non-negative and $\delta = 0$ when p_0 is not so constrained so that expected social welfare W^{sw} in (9) can be written

$$\begin{aligned}
W^{sw} &= \int_{S_l(P)} [b_l(s) - c_l(s) - \alpha p_l] f(s) ds + \int_{S_h(P)} [b_h(s) - c_h(s) - \alpha p_h] f(s) ds \\
&\quad + (\alpha - \delta) \left\{ \int_{S_l(P)} [p_l - c_l(s)] f(s) ds + \int_{S_h(P)} [p_h - c_h(s)] f(s) ds \right\} \\
&= \int_{S_l(P)} [b_l(s) - (1 + \alpha - \delta) c_l(s) - \delta p_l] f(s) ds \\
&\quad + \int_{S_h(P)} [b_h(s) - (1 + \alpha - \delta) c_h(s) - \delta p_h] f(s) ds.
\end{aligned} \tag{10}$$

The objectives (7) of a profit-maximising purchaser (for whom $\alpha = 0$) and (10) of a social-

welfare-maximising purchaser are both encompassed by the formulation

$$W = \int_{S_l(P)} [b_l(s) - (1 + \alpha - \delta) c_l(s) - \delta p_l] f(s) ds + \int_{S_h(P)} [b_h(s) - (1 + \alpha - \delta) c_h(s) - \delta p_h] f(s) ds, \quad (11)$$

where :

- $\delta = 0$ for either type of purchaser *without* constraint $p_0 \geq 0$;
- $\delta = \alpha$ for a social-welfare-maximising purchaser *with* constraint $p_0 \geq 0$;
- $\delta = 1$ for a profit-maximising purchaser *with* constraint $p_0 \geq 0$.

Cases with $\delta = 0$ correspond to the supplier's participation constraint (5) binding, so the supplier receives no informational rent from provision. In these cases, prices can be set solely to optimise provision of the two varieties to different types of recipients. In cases with $\delta > 0$, the supplier receives an informational rent corresponding to minus the terms multiplied by δ , so optimal prices need also to take account of rent extraction from the supplier.

A useful benchmark is the best outcome the purchaser could achieve if able to monitor s costlessly. Efficient provision, as defined by (1)–(3), could then be a condition of payment. Expected profit, and with $\alpha > 0$ also expected social welfare, is maximised by paying the supplier $c_i(s)$ for providing variety i to a recipient of type s . Then both expected profit W^π in (6) and expected social welfare W^{sw} in (9) are given by

$$\int_{\underline{s}}^{s'} [b_l(s) - (1 + \alpha) c_l(s)] f(s) ds + \int_{s'}^{s''} [b_h(s) - (1 + \alpha) c_h(s)] f(s) ds. \quad (12)$$

The characteristics of optimal provision depend critically on two properties of the cost functions $c_i(s)$. The first critical property is whether cost of provision is increasing or decreasing in s . If cost is increasing in s , an appropriate choice of prices will induce the supplier to provide for low s types without providing for higher s types. Thus it is feasible, though not necessarily optimal, to exclude from provision all types $s \in (s'', \bar{s}]$ that it is efficient to exclude while ensuring provision for all types $s \in [\underline{s}, s'')$. In contrast, if cost is decreasing in s , higher s types will always receive provision, whether or not that is efficient, if prices are high enough to induce the supplier to provide for lower s types. Both increasing and decreasing costs are considered here. The second critical property is whether the difference in cost between varieties h and l is decreasing or increasing in s . If this difference is decreasing, variety l is relatively less costly to provide than variety h for lower s so appropriate choice of prices will induce the supplier to provide l for lower s and h for higher s . Thus it is feasible, though again not necessarily optimal, to ensure that types lower than s' are provided with the variety l that is efficient for them while ensuring that any higher types who receive provision receive variety h . If the difference is increasing, efficient partition between the varieties provided is simply not feasible. To ensure that the cost functions have the single-crossing property, the following assumption is used for

		$c'_l(s), c'_h(s) > 0$	$c'_l(s), c'_h(s) < 0$
$c'_h(s) < c'_l(s)$	$\delta = 0$	l to $s \in [\underline{s}, s')$, h to $s \in (s', s'']$;	<i>either</i> no provision <i>or</i> l to $s \in [\underline{s}, s')$, h to $s \in (s', \bar{s}]$;
	$\delta > 0$	l to $s \in [\underline{s}, \hat{s}^*)$, possibly h to $s \in (\hat{s}^*, s_h^*]$, where $\hat{s}^* < s', s_h^* < s''$;	if l and h : l to $s \in [s_l^*, \hat{s}^*)$, h to $s \in (\hat{s}^*, \bar{s}]$, where $\underline{s} \leq s_l^* < \hat{s}^*, \hat{s}^* > s'$;
$c'_h(s) > c'_l(s)$	$\delta = 0$	either l or h to $s \in [\underline{s}, s_i^*]$, where $s_i^* \leq s''$, for $i = l, h$;	at most one of l and h to $s \in [s_i^*, \bar{s}]$, $s_i^* < s''$, for $i = l, h$;
	$\delta > 0$	if l and h : h to $s \in [\underline{s}, \hat{s}^*)$, l to $s \in (\hat{s}^*, s_l^*]$, where $s' < \hat{s}^* < s_l^* < s''$;	if l and h : h to $s \in [s_h^*, \hat{s}^*)$, l to $s \in (\hat{s}^*, \bar{s}]$, where $s_h^* < \hat{s}^* < s'$.

Table 1: Optimal provision

most of the results below.

Assumption 1 $c_h(s) - c_l(s)$ is either strictly decreasing for all $s \in [\underline{s}, \bar{s}]$ or strictly increasing for all $s \in [\underline{s}, \bar{s}]$.

The main results derived below are summarised in Table 1. The rows classify cases by whether the difference in cost $c_h(s) - c_l(s)$ between the two varieties of provision is decreasing or increasing and also by whether the supplier receives an informational rent ($\delta > 0$) or not ($\delta = 0$). The columns classify cases by whether the costs of provision $c_l(s)$ and $c_h(s)$ are increasing or decreasing. This table serves as a guide to the propositions that follow. Examples of medical care that fall into each of the four categories in the table are given in Section 5, other applications in Section 6. These can be read in conjunction with Table 1 without recourse to the detailed arguments in Sections 3 and 4.

3 Optimal prices with cost non-decreasing in s

When costs are strictly increasing in s , a supplier faced with fixed prices always supplies to lower s types if supplying to higher s types. The first part of this section analyses that case. It corresponds to the left-hand column in Table 1.

Assumption 2 $c_h(s)$ and $c_l(s)$ are strictly increasing for all $s \in [\underline{s}, \bar{s}]$.

Suppose the purchaser were to set only a single price p for provision of either variety l or h . A supplier providing to type s will then always supply whichever variety costs less, thus

incurring cost $c(s)$ given by

$$c(s) = \min [c_l(s), c_h(s)], \text{ for all } s \in [\underline{s}, \bar{s}], \quad (13)$$

which is increasing for all s and differentiable everywhere except possibly at $s = s^0$ defined by $c_l(s^0) = c_h(s^0)$, if such an s^0 exists. The benefit from that provision is $b(s)$ given by

$$b(s) = \begin{cases} b_l(s), & \text{if } c_l(s) < c_h(s); \\ \max [b_l(s), b_h(s)], & \text{if } c_l(s) = c_h(s); \\ b_h(s), & \text{if } c_l(s) > c_h(s), \end{cases} \text{ for all } s \in [\underline{s}, \bar{s}], \quad (14)$$

which is also differentiable everywhere except possibly at $s = s^0$ defined by $c_l(s^0) = c_h(s^0)$.

With the same price p for both varieties and $c'_i(s) > 0$, the supplier provides some variety for all recipient types s for which $c(s) \leq p$. Thus, if the purchaser sets price $p = c(\tilde{s})$, all types $s \leq \tilde{s}$ receive some provision. From (11), the objective function is then given by

$$\begin{aligned} \int_{\underline{s}}^{\tilde{s}} [b(s) - (1 + \alpha - \delta)c(s) - \delta c(\tilde{s})] f(s) ds \\ = \int_{\underline{s}}^{\tilde{s}} [b(s) - (1 + \alpha - \delta)c(s)] f(s) ds - \delta c(\tilde{s}) F(\tilde{s}). \end{aligned} \quad (15)$$

An optimal price is thus $c(\tilde{s}^*)$ for \tilde{s}^* that maximises the expression in (15). The first proposition characterises \tilde{s}^* . All proofs are in Appendix B.

Proposition 1 *Suppose Assumptions 1 and 2 hold and the purchaser sets a single price for provision of both varieties. Then an optimal price p satisfies $p = c(\tilde{s}^*)$, where $\tilde{s}^* = \underline{s}$, or $\tilde{s}^* = s^0$ uniquely defined by $c_l(s^0) = c_h(s^0)$, or \tilde{s}^* satisfies*

$$b(\tilde{s}^*) - (1 + \alpha)c(\tilde{s}^*) = \delta c'(\tilde{s}^*) \frac{F(\tilde{s}^*)}{f(\tilde{s}^*)}. \quad (16)$$

Moreover, $\tilde{s}^* \leq s''$, with strict inequality for $\delta > 0$, and, if $c_l(\underline{s}) < c_h(\underline{s})$, $\tilde{s}^* > \underline{s}$.

This result illustrates two fundamental points. First, while by manipulating the price the purchaser can determine how far up the interval from \underline{s} to \bar{s} the cutoff point \tilde{s} for provision is, when $\delta > 0$ it is never worth raising that cutoff as far as s'' , at which all those for whom some variety is efficient receive provision. The reason is that raising the price to induce the supplier to provide for higher s types increases the informational rent on provision for infra-marginal types given by the right-hand side of (16). Second, with just one price, the supplier provides any type s receiving provision with whichever variety costs less. Thus if, for example, $c_h(s) < c_l(s)$ for all s , the supplier provides only variety h and any types $s < s'$ receive that variety even though variety l is efficient for them. By setting different prices for the two varieties, the purchaser can influence both the types that receive each variety and the informational rent to the supplier.

It is straightforward to specify conditions under which (16) has a unique solution. One set of conditions is $b_i(s) - (1 + \alpha) c_i(s)$ decreasing for $i \in \{l, h\}$, $c_i(s)$ convex for $i \in \{l, h\}$, and $F(s)$ log concave (that is, $F(s)/f(s)$ non-decreasing) for all s . The last of these is commonly assumed in the literature on contracting under asymmetric information (see Laffont and Tirole (1993)) and is satisfied by such standard distributions as the normal, the uniform and the chi-squared. Under those additional assumptions, the left-hand side of (16) is decreasing and the right-hand side non-decreasing, so there can be at most one solution. Uniqueness is not, however, required for the results derived here.

For different prices set by the purchaser, consider Figure 1 in which $c_j(s)$ and $c_k(s)$ are the increasing cost functions and j may correspond to either l or h . Suppose the purchaser sets prices p_j and p_k . For $s > s_k$, cost is greater than price for both varieties, so the supplier provides neither. For $s < \hat{s}$, the difference between price and cost is greater with variety j , so the supplier provides j . For $s \in (\hat{s}, s_k)$, that difference is greater for variety k , so the supplier provides k . Formally, define s_i for $i \in \{l, h\}$ by

$$s_i = \begin{cases} s \text{ such that } p_i - c_i(s) = 0, \text{ if there exists such an } s \in [\underline{s}, \bar{s}]; \\ \bar{s}, \text{ if } p_i - c_i(s) > 0 \text{ for all } s \in [\underline{s}, \bar{s}]; \\ \underline{s}, p_i - c_i(s) < 0 \text{ for all } s \in [\underline{s}, \bar{s}]; \end{cases}, i \in \{l, h\}. \quad (17)$$

In words, s_i is the highest s to whom the supplier is willing to provide variety i given p_i . Define $k \in \{l, h\}$ by

$$s_k = \max(s_l, s_h). \quad (18)$$

Thus s_k is the highest s to whom the supplier provides either variety given p_l and p_h . By Assumption 2, all of s_l , s_h and s_k are unique for given p_l , p_h . Finally, define \hat{s} by

$$\hat{s} = \begin{cases} s \text{ such that } c_h(s) - c_l(s) = p_h - p_l, \text{ if there exists such an } s \in [\underline{s}, s_k]; \\ \underline{s}, \text{ otherwise.} \end{cases} \quad (19)$$

Under Assumption 1, \hat{s} is unique for given p_l and p_h .

Consider $\hat{s} = \underline{s}$. By the definition of k in (18) and Assumption 1, for $j \in \{l, h\}$ and $j \neq k$, $p_k - c_k(s) > p_j - c_j(s)$ for $s \in (s_j, s_k]$. With $c_i(s)$ continuous for $i \in \{l, h\}$, this must also be true for $s \in [\underline{s}, s_j]$ if there is no s satisfying the top line of (19). Thus, if $\hat{s} = \underline{s}$, the supplier provides variety k to any s receiving provision. If $\hat{s} > \underline{s}$, the supplier provides variety j to recipient types $s \in [\underline{s}, \hat{s})$, variety k to types $s \in (\hat{s}, s_k]$, and neither variety to $s \in (s_k, \bar{s}]$. It is indifferent to which variety is provided for type $s = \hat{s}$. Given this decision rule of the supplier, the objective function in (11) can be written

$$W = \int_{\underline{s}}^{\hat{s}} [b_j(s) - (1 + \alpha - \delta) c_j(s)] f(s) ds + \int_{\hat{s}}^{s_k} [b_k(s) - (1 + \alpha - \delta) c_k(s)] f(s) ds - \delta [F(s_k) p_k - F(\hat{s}) (p_k - p_j)]. \quad (20)$$

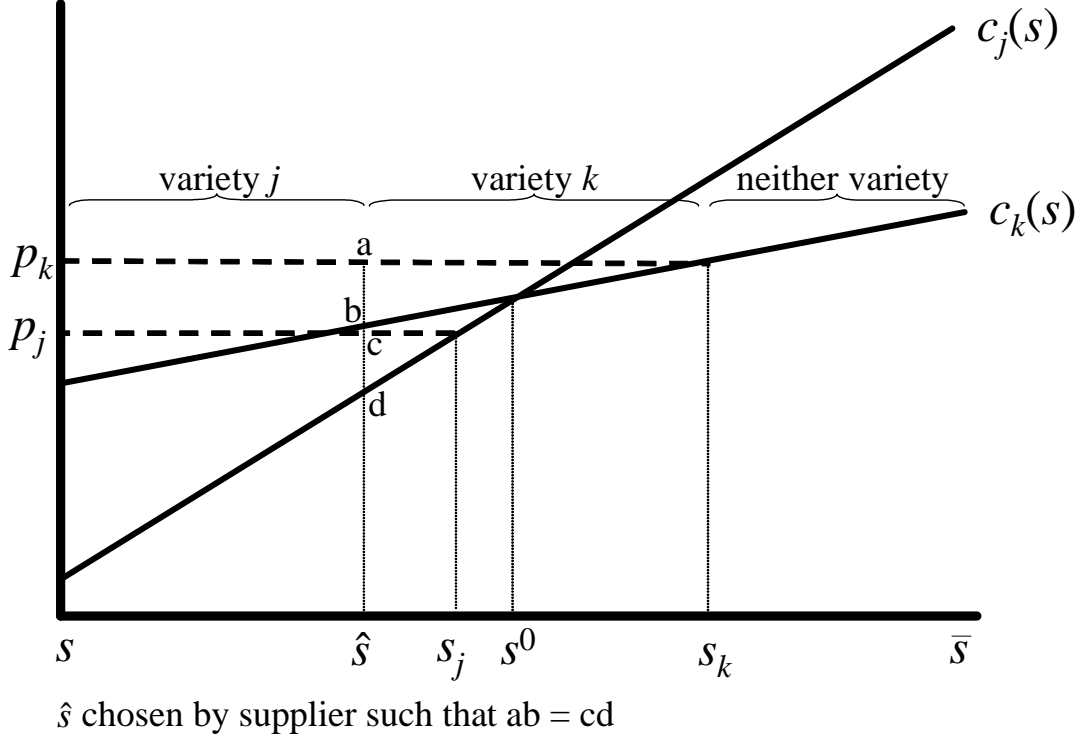


Figure 1: Supplier provision decisions with increasing costs

In view of (19), p_j can be replaced by $p_k - c_k(\hat{s}) + c_j(\hat{s})$ without loss of generality because, if the top line of (19) applies, these are equal and, if the bottom line of (19) applies, $\hat{s} = \underline{s}$ so $F(\hat{s}) = 0$ and p_j does not enter the maximand (20). Moreover, given (17) and (18), p_k can be replaced by $c_k(s_k)$. With these substitutions, the purchaser's optimization problem can be expressed in terms of \hat{s} and s_k as

$$\begin{aligned} \max_{\hat{s}, s_k} \int_{\underline{s}}^{\hat{s}} [b_j(s) - (1 + \alpha - \delta) c_j(s)] f(s) ds + \int_{\hat{s}}^{s_k} [b_k(s) - (1 + \alpha - \delta) c_k(s)] f(s) ds \\ - \delta \{ F(s_k) c_k(s_k) - F(\hat{s}) [c_k(\hat{s}) - c_j(\hat{s})] \}. \end{aligned} \quad (21)$$

First-order necessary conditions that must be satisfied by optimal \hat{s} and s_k interior to $[\underline{s}, \bar{s}]$, denoted \hat{s}^* and s_k^* respectively, are

$$\begin{aligned} [b_j(\hat{s}^*) - (1 + \alpha - \delta) c_j(\hat{s}^*)] f(\hat{s}^*) - [b_k(\hat{s}^*) - (1 + \alpha - \delta) c_k(\hat{s}^*)] f(\hat{s}^*) \\ + \delta \{ f(\hat{s}^*) [c_k(\hat{s}^*) - c_j(\hat{s}^*)] + F(\hat{s}^*) [c'_k(\hat{s}^*) - c'_j(\hat{s}^*)] \} = 0 \\ [b_k(s_k^*) - (1 + \alpha - \delta) c_k(s_k^*)] f(s_k^*) - \delta [f(s_k^*) c_k(s_k^*) + F(s_k^*) c'_k(s_k^*)] = 0. \end{aligned}$$

An interior solution has $\hat{s}^*, s_k^* \in (\underline{s}, \bar{s})$ and hence $f(\hat{s}^*), f(s_k^*) > 0$, so these necessary conditions can be simplified to

$$[b_j(\hat{s}^*) - b_k(\hat{s}^*)] - (1 + \alpha)[c_j(\hat{s}^*) - c_k(\hat{s}^*)] = \delta [c'_j(\hat{s}^*) - c'_k(\hat{s}^*)] \frac{F(\hat{s}^*)}{f(\hat{s}^*)} \quad (22)$$

$$b_k(s_k^*) - (1 + \alpha)c_k(s_k^*) = \delta c'_k(s_k^*) \frac{F(s_k^*)}{f(s_k^*)}. \quad (23)$$

Lemma 1 *Suppose Assumptions 1 and 2 hold. Then, for optimal prices, there exist $s_k^* \in (\underline{s}, s'']$ satisfying (23) and $\hat{s}^* \in [\underline{s}, s_k^*]$ such that the supplier provides recipient types $s \in [\underline{s}, \hat{s}^*)$ with variety $j \neq k$, recipient types $s \in (\hat{s}^*, s_k^*)$ with variety k , and recipient types $s \in (s_k^*, \bar{s}]$ with neither variety. If $\hat{s}^* > \underline{s}$, \hat{s}^* satisfies (22). If $\hat{s}^* = \underline{s}$, the left-hand side of (22) must be non-positive. If $\delta > 0$, $s_k^* < s''$.*

Lemma 1 establishes that the supplier provides one variety j to recipient types below \hat{s}^* , the other variety k to recipient types between \hat{s}^* and s_k^* , and neither variety to recipient types above s_k^* . It does not, however, specify whether j corresponds to l or to h — that depends on whether $c_h(s) - c_l(s)$ is increasing or decreasing. Consider first the case $c_h(s) - c_l(s)$ strictly decreasing for all $s \in [\underline{s}, \bar{s}]$.

Proposition 2 *Suppose Assumption 2 holds and $c_h(s) - c_l(s)$ is strictly decreasing for all $s \in [\underline{s}, \bar{s}]$.*

1. *It is always strictly optimal to ensure variety l is provided.*
2. *It is strictly optimal to ensure variety h is provided if*

$$\frac{b_h(s) - (1 + \alpha)c_h(s)}{c'_h(s)} > \frac{b_l(s) - (1 + \alpha)c_l(s)}{c'_l(s)} \quad (24)$$

for $s = s_l^$ defined by (23) with $k = l$. Condition (24) holds for all $s \in (s', \bar{s}]$.*

3. *If it is optimal to have both varieties provided: (a) variety j of Lemma 1 corresponds to l and variety k to h ; (b) $\hat{s}^* \leq s'$ and $s_h^* \leq s''$, with the inequalities strict for $\delta > 0$; and (c) optimal prices p_i^* for $i \in \{l, h\}$ satisfy $p_l^* > p_h^*$ if $c_l(\underline{s}) > c_h(\underline{s})$ and $p_l^* < p_h^*$ if $c_l(s') < c_h(s')$, or $c_l(s') \leq c_h(s')$ in the case $\delta > 0$.*
4. *For $\delta = 0$, both varieties are provided, $\hat{s}^* = s'$ and $s_h^* = s''$.*

Proposition 2 corresponds to the upper left box in Table 1. The key to understanding it is that, with $c_h(s) - c_l(s)$ strictly decreasing, the relative cost of providing variety l is lower for lower s types and l corresponds to j in Figure 1. Since it is low s types for which variety l is efficient, with appropriate prices the supplier could be induced to provide each variety to those, and only those, for whom that variety is efficient. For $\delta = 0$ (that is, upfront payments ensure

the supplier expects no rent from provision), that is optimal, as Part 4 of Proposition 2 shows. For $\delta > 0$, however, that is not optimal because of the informational rent the supplier would receive. Providing for more types requires an increase in price to cover the higher cost. But that increases the rent to the supplier on all infra-marginal types in exactly the same way as discussed in connection with Proposition 1. The essential point can be seen from (23). The left-hand side corresponds to the net benefit of providing variety k to the marginal type s_k^* . The right-hand side, which is strictly positive for $s_k^* > \underline{s}$ when $c'_k(s) > 0$ and $\delta > 0$, corresponds to the deadweight loss from the increased rent the supplier receives on all $s < s_k^*$. Because of that increased rent, it is optimal to stop short of having provision made for those for whom the net benefit of provision, while still positive, is sufficiently low. But, as shown in Part 1 of Proposition 2, it is always optimal to ensure that variety l , at least, is provided because for type \underline{s} there are no infra-marginal types being treated, so it is always worth ensuring that they receive l . Moreover, if for the type s_l^* that would be marginal in the event that only variety l were provided the relative benefit of providing variety h is sufficiently large, then it is optimal also to have variety h provided. That is the content of Part 2 of Proposition 2.

If both varieties are provided when $\delta > 0$, it is not optimal to have variety h provided for those for whom the net benefit is sufficiently small because of the increased rent the supplier would receive, so $s_h^* < s''$. But it is also, for a similar reason, optimal to have some for whom variety l is efficient receive variety h . That can be seen from (22). With $j = l$ and $k = h$, the left-hand side is the difference in the net benefit of type \hat{s}^* receiving variety l over receiving variety h . The right-hand side is the difference in the informational rent from inducing the supplier to provide type \hat{s}^* with variety l rather than variety h . With $c'_l(s) > c'_h(s)$, the additional rent from having type \hat{s}^* supplied with l is greater than from having that type supplied with h . At an optimum, this additional rent just counterbalances the reduced net gain from having type \hat{s}^* supplied with h . This means setting $\hat{s}^* < s'$. The results on relative prices when both varieties are provided follow directly from (19), which implies $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. The conditions given in Part 3 of Proposition 2 are sufficient to determine the sign of $c_l(\hat{s}^*) - c_h(\hat{s}^*)$ given $\hat{s}^* < s'$.

A number of conclusions for when it is preferable to use two different prices follow directly from Proposition 2. First, if a single price would result in only variety h being provided (because, for example, $c_h(s) < c_l(s)$ for all s), it follows from Part 1 that there is a strict gain from having two different prices. Second, if a single price would result in only variety l being provided (because, for example, $c_l(s) < c_h(s)$ for all s) and $\tilde{s}^* > s'$, it follows from Part 2 that there is a strict gain from having two different prices whenever (24) is satisfied. Third, Part 3 gives conditions for the two prices to be different. A particular case in which one of those conditions holds is that studied by McClellan (1996b) and widely assumed to apply to hospital treatments, namely the treatment appropriate for low cost cases is less costly for all recipients (though less beneficial for at least some). Then $c_l(s) < c_h(s)$ for all s , and in particular for $s = s'$, so $p_l^* < p_h^*$. Another case is that in which the benefits of the two varieties are the same ($b_h(s) = b_l(s)$ for all s). It is then a direct consequence of (1) and (2) that $c_l(s') = c_h(s')$, so

again $p_l^* < p_h^*$ provided $\delta > 0$. Finally, even if it is optimal to have only one variety provided, it may still be necessary to have two different prices. If, for example, it is optimal to have only variety l provided but $c_l(s) > c_h(s)$ for all s , the supplier would provide only variety h if the prices for the two were the same.

Now consider the case $c_h(s) - c_l(s)$ strictly increasing for all $s \in [\underline{s}, \bar{s}]$.

Proposition 3 *Suppose Assumption 2 holds and $c_h(s) - c_l(s)$ is strictly increasing for all $s \in [\underline{s}, \bar{s}]$.*

1. *Either $\hat{s}^* > s'$ or $\hat{s}^* = \underline{s}$.*
2. *For $\delta = 0$, only one variety $i \in \{l, h\}$ is provided, with $s_i^* \leq s''$.*
3. *If it is optimal to have both varieties provided: (a) variety j of Lemma 1 corresponds to h and variety k corresponds to l ; and (b) optimal prices p_i^* for $i \in \{l, h\}$ satisfy $p_l^* < p_h^*$ if $c_l(s') \leq c_h(s')$ and $p_l^* > p_h^*$ if $c_l(s'') \geq c_h(s'')$.*

The results in Proposition 3 correspond to the lower left box in Table 1. With $c_h(s) - c_l(s)$ strictly increasing, variety h corresponds to j in Figure 1. While, as with $c_h(s) - c_l(s)$ strictly decreasing in Proposition 2, it is still always optimal to ensure that at least one variety is supplied (that follows from $s_k > \underline{s}$ in Lemma 1), it is no longer necessarily variety l . The reason is that, when $c_h(s) - c_l(s)$ is strictly increasing, the relative cost of providing l rather than h is higher for low s types for whom variety l is efficient. Thus, to ensure that low s types receive variety l , p_h will have to be set sufficiently low that nobody receives variety h .

For $\delta = 0$ (that is, the supplier receives no informational rent) the implication is that only one variety is provided. If both were provided, some for whom variety l is efficient would receive variety h and some for whom variety h is efficient variety l . So, unlike the case $\delta = 0$ under the conditions of Proposition 2, the efficient outcome is not feasible. If instead only one variety is provided, one of those inefficiencies is removed and, with no concern for rent extraction from the supplier, that increases the purchaser's payoff. Which variety it is optimal to provide depends on whether the loss in payoff is less from providing types $s < s'$ with variety h or from providing types $s \in [s', s_l^*]$ with variety l .

For $\delta > 0$, there is the additional issue of extracting rent from the supplier. If it is optimal to provide both varieties in order to do this, Part 1 of Proposition 3 implies that $\hat{s}^* > s'$. The implication is that not only do all those with $s \leq s'$ for whom variety l is efficient receive variety h , but also some types $s > s'$ for whom variety h is efficient. The reason is that, if $\hat{s}^* < s'$, no types for whom variety h is efficient actually receive it, so it would be better not to have variety h supplied at all. But providing variety l , even if only to some types $s > s'$ for whom variety h is efficient, may reduce the rent to the supplier. With $c_h(s) - c_l(s)$ strictly increasing, and hence $c'_h(s) > c'_l(s)$, the cost for variety l rises less steeply as s increases than the cost for variety h . So providing variety l to additional types at the margin requires less increase in price, and hence less additional rent to the supplier on infra-marginal types, than

providing variety h . That is reflected in the first-order condition (23). With $c'_h(s) > c'_l(s)$, the right-hand side of (23) is smaller for given s if $k = l$ than if $k = h$, so it may be that more types receive provision of some variety when variety l is provided even when the left-hand side is larger for $k = h$. Again, the results on relative prices when both varieties are provided follow directly from (19), which implies $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. The conditions given in Part 3 of Proposition 3 are sufficient to determine the sign of $c_l(\hat{s}^*) - c_h(\hat{s}^*)$ given that $\hat{s}^* > s'$.

The case with non-decreasing cost not covered by Propositions 2 and 3 is that in which the difference between $c_h(s)$ and $c_l(s)$ is the same for all s , so $c'_h(s) = c'_l(s)$ for all s . Then, unless prices are set such that $p_l - c_l(s) = p_h - c_h(s)$, only one of the varieties will be supplied. So even a small mistake would have a big effect on the provision made. Moreover, the effect on the purchaser's payoff need not be small because the “wrong” variety is provided to infra-marginal types for whom the difference in benefit between the two varieties may be large. Thus, even if the purchaser is only slightly uncertain about the supplier's costs for the two varieties, the implications could be major.

Provided the purchaser sets the prices exactly right, the supplier is indifferent as to which variety is supplied for any s when $c'_h(s) = c'_l(s)$. The supplier would then be willing to provide each s with whichever variety the purchaser would prefer. Since, the right-hand side of (22) is zero in this case, it follows from that condition and from (1) and (2) that the supplier would wish to have $\hat{s}^* = s'$, with variety l provided to $s < s'$ and variety h to all types $s > s'$ who receive provision. Then, all those types who receive provision receive the variety that is efficient for them. But, for $c'_h(s) > 0$ and $\delta > 0$, the right-hand side of (23) is strictly positive, so it follows from (2) that it is not optimal to provide for all those types for whom provision would be efficient. Only in the cases $c'_h(s) = c'_l(s) = 0$ or $\delta = 0$ is some provision supplied for all those for whom it is efficient. If the costs of the two varieties are actually the same, then a single price ($p_l = p_h$) is optimal. The supplier is, however, equally willing to supply not what the *purchaser* wants but what the *recipient* wants, a possibility that seems plausible in the context of provision of, in particular, medical care. It is assumed in much of the literature that physicians are influenced by what their patients want even if this is not what the purchaser would want—see the survey by McGuire (2000). In that case, it is potentially disastrous to leave the supplier indifferent for each s as to whether provision is made and, if it is, which variety of provision. The supplier will then provide each s with whichever variety has higher benefit to the patient and will provide some variety as long as the higher of the two benefits is positive. Thus decisions on which variety to provide will be made entirely on the basis of benefit to the patient, without consideration of cost. Only in special cases, for example, the two varieties cost the same and $s'' = \bar{s}$ so that it is efficient to provide for all types, will this result in efficient decisions about the variety to be provided.

4 Optimal prices with cost decreasing in s

The previous section analysed optimal prices with costs of provision non-decreasing in s . This section analyses optimal prices with costs of provision decreasing in s , corresponding to the right-hand column in Table 1.

Assumption 3 $c_h(s)$ and $c_l(s)$ are strictly decreasing for all $s \in [\underline{s}, \bar{s}]$.

The implication is that potential recipients for whom neither variety of provision is efficient are the cheapest to supply instead of the most expensive. That is the only change in assumption from the previous section. In particular, the specification of which variety of provision is efficient in (1)-(3) is retained. As before, both $c_h(s) - c_l(s)$ decreasing and $c_h(s) - c_l(s)$ increasing are considered. The analysis of both these cases closely mirrors that of the preceding section, so the arguments are presented only briefly.

The analysis with Assumption 3 is illustrated in Figure 2, which differs from Figure 1 in that costs are now decreasing. Thus, if the supplier's payoff from providing type s with variety i is positive, the payoff from providing types higher than s with i is also positive. Retain the definition of s_i in (17) but change the definition of k from that in (18) to

$$s_k = \min(s_l, s_h). \quad (25)$$

This is the lowest s to whom the supplier provides either variety given p_l and p_h . Also, change the definition of \hat{s} from that in (19) to

$$\hat{s} = \begin{cases} s \text{ such that } c_h(s) - c_l(s) = p_h - p_l, & \text{if there exists such an } s \in [s_k, \bar{s}]; \\ \bar{s}, & \text{otherwise.} \end{cases} \quad (26)$$

Under Assumption 1, \hat{s} is unique for given p_l and p_h . Figure 2 illustrates these definitions.

Consider $\hat{s} = \bar{s}$. By the definition of k in (25) and Assumption 1, for $j \in \{l, k\}$ and $j \neq k$, $p_k - c_k(s) > p_j - c_j(s)$ for $s \in (s_k, s_j]$ and, with $c_i(s)$ continuous for $i \in \{l, h\}$, this must also be true for $s \in [s_j, \bar{s}]$ if there is no s satisfying the top line of (26). Thus, if $\hat{s} = \bar{s}$, the supplier provides variety k to any s receiving provision. If $\hat{s} < \bar{s}$, the supplier provides variety k to recipient types $s \in [s_k, \hat{s})$, variety j to types $s \in (\hat{s}, \bar{s}]$, and neither variety to types $s \in [\underline{s}, s_k)$. It is indifferent to which variety is provided to type $s = \hat{s}$. Given this decision rule of the supplier, the purchaser's objective function in (11) can be written

$$W = \int_{s_k}^{\hat{s}} [b_k(s) - (1 + \alpha - \delta) c_k(s)] f(s) ds + \int_{\hat{s}}^{\bar{s}} [b_j(s) - (1 + \alpha - \delta) c_j(s)] f(s) ds - \delta \{ [1 - F(s_k)] p_k - [1 - F(\hat{s})] (p_k - p_j) \}. \quad (27)$$

In view of (26), p_j can be replaced by $p_k - c_k(\hat{s}) + c_j(\hat{s})$ without loss of generality because, if the top line of (26) applies, these are equal and, if the bottom line of (26) applies, $\hat{s} = \bar{s}$ so

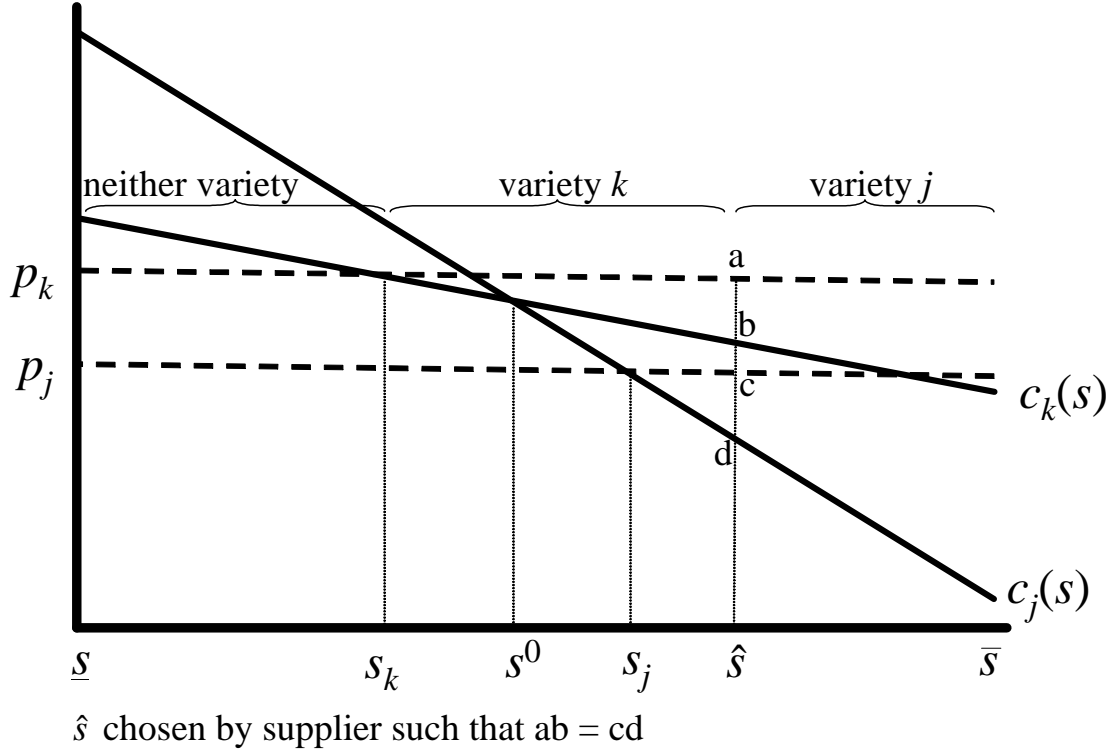


Figure 2: Supplier provision decisions with decreasing costs

$1 - F(\hat{s}) = 0$ and p_j does not enter the maximand (27). Moreover, given (17) and (25), p_k can be replaced by $c_k(s_k)$. With these substitutions, the purchaser's optimization problem can be expressed in terms of \hat{s} and s_k as

$$\max_{\hat{s}, s_k} \int_{s_k}^{\hat{s}} [b_k(s) - (1 + \alpha - \delta) c_k(s)] f(s) ds + \int_{\hat{s}}^{\bar{s}} [b_j(s) - (1 + \alpha - \delta) c_j(s)] f(s) ds - \delta \{ [1 - F(s_k)] c_k(s_k) - [1 - F(\hat{s})] [c_k(\hat{s}) - c_j(\hat{s})] \}. \quad (28)$$

The first-order necessary conditions (22) and (23) for an interior solution are then replaced by

$$[b_k(\hat{s}^*) - b_j(\hat{s}^*)] - (1 + \alpha) [c_k(\hat{s}^*) - c_j(\hat{s}^*)] = -\delta [c'_k(\hat{s}^*) - c'_j(\hat{s}^*)] \frac{1 - F(\hat{s}^*)}{f(\hat{s}^*)} \quad (29)$$

$$b_k(s_k^*) - (1 + \alpha) c_k(s_k^*) = -\delta c'_k(s_k^*) \frac{1 - F(s_k^*)}{f(s_k^*)}. \quad (30)$$

One difference this case makes to the results of Lemma 1 is that it may now be optimal not to provide either variety to any $s \in [s, \bar{s}]$. The reason is that, with cost decreasing in s , the supplier will always provide for types $s \in (s'', \bar{s}]$ if any $s \leq s''$ receive provision. But by (3) it is efficient to provide for only those types $s \leq s''$ and the efficiency loss from providing for $s \in (s'', \bar{s}]$ may more than outweigh the gain from providing for some $s < s''$. This applies even if the supplier receives no informational rent ($\delta = 0$). It can, of course, apply only if

$s'' < \bar{s}$ because otherwise it is efficient to have provision for all types. Another difference is that it may be optimal to have all types $s \in [\underline{s}, \bar{s}]$ receive provision if $b_k(\underline{s}) - (1 + \alpha) c_k(\underline{s})$ is sufficiently large. Thus s_k^* may not be interior to $[\underline{s}, \bar{s}]$ and may, indeed, be at *either* endpoint of this interval. The main results for cost decreasing in s corresponding to those in Propositions 2 and 3 for cost increasing in s can be summarised as follows.

Proposition 4 *Suppose Assumption 3 holds.*

1. *Suppose $c_h(s) - c_l(s)$ is strictly decreasing for all $s \in [\underline{s}, \bar{s}]$. If it is optimal to have both varieties provided: (a) variety k corresponds to l and variety j corresponds to h ; (b) $\hat{s}^* > s'$ for $\delta > 0$ and $\hat{s}^* = s'$ for $\delta = 0$; and (c) optimal prices p_i^* for $i \in \{l, h\}$ satisfy $p_l^* < p_h^*$ if $c_l(\bar{s}) < c_h(\bar{s})$ and $p_l^* > p_h^*$ if $c_l(s') > c_h(s')$, or $c_l(s') \geq c_h(s')$ in the case $\delta > 0$.*
2. *Suppose $c_h(s) - c_l(s)$ is strictly increasing for all $s \in [\underline{s}, \bar{s}]$. For $\delta = 0$, it is optimal to have at most one variety provided. If for $\delta > 0$ it is optimal to have both varieties provided: (a) variety j corresponds to l and variety k corresponds to h ; (b) $\hat{s}^* < s'$; and (c) optimal prices p_i^* for $i \in \{l, h\}$ satisfy $p_l^* < p_h^*$ if $c_l(\underline{s}) \leq c_h(\underline{s})$ and $p_l^* > p_h^*$ if $c_l(s') \geq c_h(s')$.*

Part 1 of Proposition 4 corresponds to the upper right box in Table 1. It is akin to Proposition 2 in that the cost of providing variety l relative to variety h is lowest for low s and highest for high s , in line with what it is efficient. Thus, with prices fixed appropriately, the supplier could be induced to provide variety l only to those for whom it is efficient. As with Proposition 2, that is optimal only when the supplier receives no informational rent ($\delta = 0$). Otherwise, to keep down that rent, it is optimal to set prices such that $\hat{s}^* > s'$, so some types for whom variety h is efficient receive variety l . But, unlike in Proposition 2, there is always a problem with types for whom neither variety is efficient if there are any (that is, if $s'' < \bar{s}$). With cost decreasing in s , these always receive some provision at any prices that will induce provision to any types for whom provision is efficient.

Part 2 of Proposition 4 corresponds to the lower right box in Table 1. It is akin to Proposition 3 in that the cost of providing variety l relative to variety h is lowest for high s and highest for low s , the opposite of what it is efficient. Thus no prices will induce an efficient choice of cutoff between varieties l and h . For this reason, when the supplier receives no rent ($\delta = 0$), it is optimal to have at most one variety provided — any other choice would have some types for whom variety l is efficient receive variety h and some for whom variety h is efficient receive variety l , one of which can be prevented if only one variety is provided. When $\delta > 0$, there is the additional issue of reducing the supplier's rent. If it is optimal to provide both varieties in order to do this then, with $\hat{s}^* < s'$ and j corresponding to l , all those for whom variety h is efficient actually receive variety l . Variety h is received only by types for whom variety l would be efficient, an outcome that can be optimal only because it enables those types to receive some variety at a lower rent to the supplier. Again, if there are types for whom neither variety is

efficient (that is, if $s'' < \bar{s}$), they still receive provision, in this case variety l which, in view of (3), is even more inefficient for them than is variety h .

Proposition 4 completes the development of the results summarised in Table 1. The remainder of the paper is concerned with applications.

5 Application to medical care

An important practical application of the analysis is to the prices set for medical services, for example prices for diagnosis-related groups (DRGs) such as those used in the US Medicare system. McClellan (1997, p. 93) comments that, with Medicare, “over 40 percent of DRGs are related not to diagnoses . . . but to the performance of specific intensive procedures” (over 220 out of 480 DRGs in 1993). Different procedures for the same basic diagnosis correspond to different varieties of provision in the model used here. Moreover, fixed payments per patient for each DRG (which correspond to prices in the model) account for the vast majority of total payments, 95% averaged over all DRGs according to McClellan (1997).² The model applies to differences in patient type for which the appropriate choice of treatment is a matter of genuine medical judgement. Costs and benefits in the model correspond to the expected costs and benefits given the information about a patient’s condition at the time of decision.

The results derived here can be applied directly to data from cost-effectiveness studies such as Patterson et al. (1995) of different test procedures for coronary artery disease (CAD). Two basic test procedures are compared there: (1) coronary angiography, a high-precision but invasive procedure with non-negligible risks, by itself; and (2) initial use of a lower precision but less costly and non-invasive procedure, followed by coronary angiography for those with positive results on the initial test. The initial tests considered are Exercise ECG (ExECG), Positron Emission Tomography (PET), and Single Photon Emission Computed Tomography (SPECT). The study derives the expected cost of each procedure and the expected benefit in terms of the gain in Quality Adjusted Life Years (QALYs) for different pre-test probabilities of CAD based on patient characteristics observable to a physician. To calculate the net benefit as defined here requires only values for QALY gain and, in the case of a social-welfare-maximising purchaser, the social cost of public funds α . Figure 3 shows the net benefit for given pre-test probabilities for a value of \$50,000 per QALY gain (a figure widely used in the literature) and $\alpha = 0$. For these values, coronary angiography by itself and PET followed by coronary angiography

²Since hospitals are required to report the costs of treatment for individual patients, the potential exists for conditioning payment on cost by using an optimal cost-sharing payment of the type discussed in Baron and Myerson (1982) and Laffont and Tirole (1993)—*supply-side cost sharing* in the terminology of Ellis and McGuire (1993). Indeed, reported costs are actually used to determine cost outlier payments for unusually expensive patients. Newhouse (1996) has put the case for using more cost-based payment for health services, Chalkley and Malcomson (2000) review the formal models, and Chalkley and Malcomson (2002) assess the cost savings that might be achieved if cost sharing were extended to all Medicare treatments. However, one of the main reasons for moving to DRG-based payments was concern about the extent to which reported costs reflect the true costs of providing the appropriate treatment for patients. Fixed payments avoid the need to monitor costs effectively for each patient. They also have the advantage of ensuring that cost-reducing effort is always efficient.

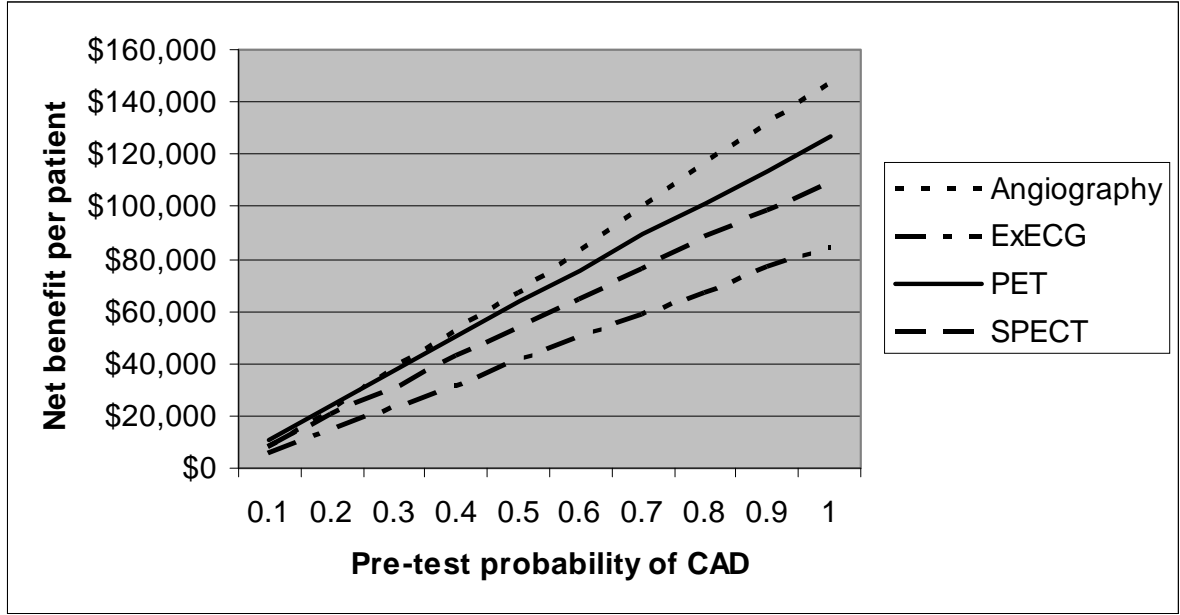


Figure 3: Net benefits of tests for coronary artery disease, \$50,000 per QALY, $\alpha = 0$, other data from Patterson et al. (1995)

between them always dominate for all pre-test probabilities, so only these two procedures need to be considered. The former has higher net benefit for probabilities above 0.25, the latter for probabilities below 0.25. These conclusions are reasonably robust to the value of a QALY. With the value halved or doubled, coronary angiography and PET still dominate and the efficient cutoff between them changes by less than 0.1.

To apply the results derived here, it is necessary to assess how pre-test probability of CAD relates to patient type s and how costs change with s . Although benefits and costs for pre-test probabilities less than 0.1 are not given in the study, benefit cannot be positive for a pre-test probability of zero and so net benefit becomes negative as the pre-test probability goes to zero. To correspond to the specification in (1)–(3), higher s must correspond to lower pre-test probability. Thus coronary angiography corresponds to variety of provision l that is efficient for low s and PET to variety h , with s' corresponding to a pre-test probability of approximately 0.25. The value of s'' at which net benefit becomes negative cannot be calculated from the data in the study but this turns out not to affect the application. Figure 4 shows the expected costs of the two dominating procedures for given pre-test probabilities. The cost for PET increases substantially with pre-test probability, primarily because there is a greater likelihood of following it with coronary angiography, that for coronary angiography does not increase significantly and is treated as constant in the study. With s measured inversely to pre-test probability, this is a case with cost decreasing in s , corresponding to the right-hand column in Table 1. (A constant cost of coronary angiography does not create a problem for this.) Moreover, with coronary angiography corresponding to l and PET to h , $c'_h(s) - c'_l(s) < 0$. Thus the relevant box in Table 1 is the upper right one. The implications are as follows. If

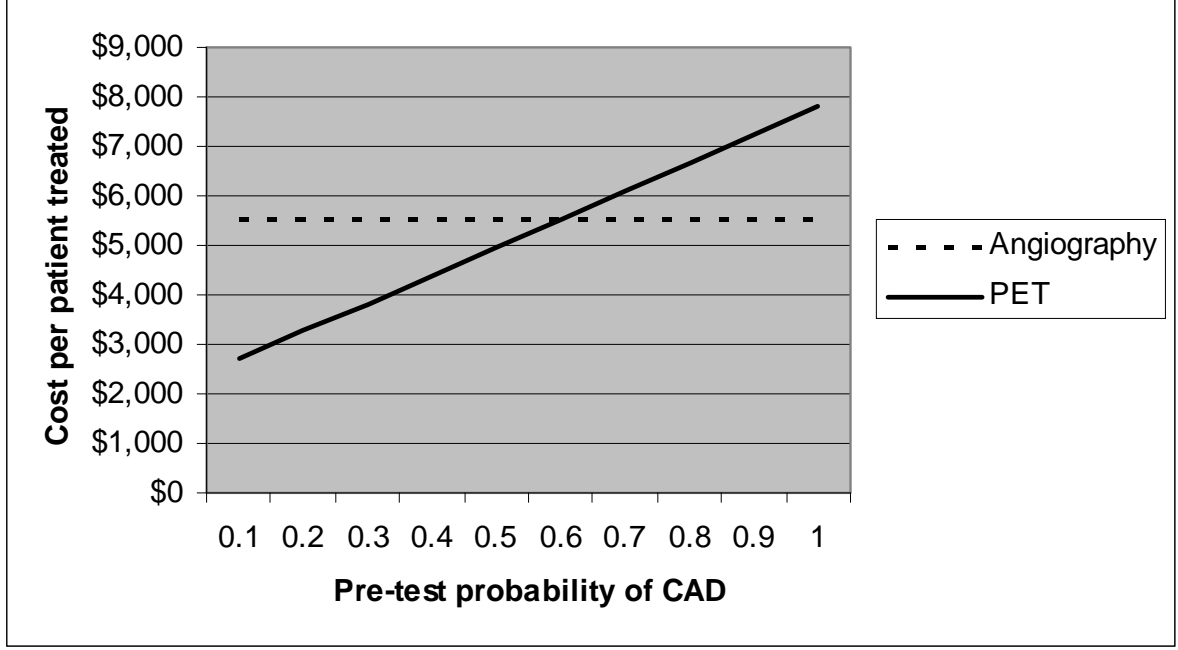


Figure 4: Costs of tests for coronary artery disease, data from Patterson et al. (1995)

prices are set high enough that the supplier is willing to provide tests to type s , it is also willing to supply to any higher types. So if any $s \leq s''$ receive a test, $s > s''$ who want a test will also receive one. If, despite this, it is worthwhile having either test procedure applied (which presumably it is given their widespread use) and if the supplier receives no ex ante rent ($\delta = 0$), it is optimal to set prices so that the cutoff between angiography and PET is at the efficient level corresponding to a pre-test likelihood of approximately 0.25. With a constant cost to coronary angiography, the optimal price is necessarily equal to that cost, so $p_l^* = \$5,500$. To induce the supplier to choose the optimal cutoff, the price of PET has, by (26), to be set so that the difference between price and cost for the two procedures is the same for $s = s'$. That gives $p_h^* \approx \$3,800$. If the supplier cannot be prevented from receiving an ex ante rent ($\delta > 0$), it is optimal to set p_h lower than that so that the cutoff is above s' (that is, at a pre-test probability of CAD below 0.25, resulting in some types for whom PET is efficient receiving coronary angiography). Precisely how much lower depends on the distribution of patient types but, however much it is, $p_l^* > p_h^*$ as in Part 1 of Proposition 4. With the cost of angiography the same for all types, there is no rent to the supplier from providing this so, with net benefit positive for all $s < s'$, it is optimal to provide the procedure for all these types.

This application illustrates an important point about the specification for which a price is paid. Since the procedure using PET consists of two separate tests (PET itself, followed eventually by coronary angiography if the result from PET is positive), it would be possible to price the two parts separately, with that for coronary angiography paid only if it is actually carried out. That would result in two procedures both with costs more or less independent of s . But that might not be in the purchaser's interest for the two reasons given in the discussion of the

constant cost case at the end of Section 3. First, if price is set equal to cost for both, the supplier may use the procedure the patient prefers rather than the one the purchaser prefers. Second, if the prices are not set precisely right, the supplier will provide only one of the procedures despite it being optimal to have both supplied. By specifying PET followed by coronary angiography as a single variety of provision with cost decreasing in s , the purchaser ensures that the supplier strictly prefers one procedure to the other for all but the cutoff type and that marginal errors in pricing result in only marginal deviations from the optimal cutoff. On the other hand, with constant costs for both procedures, the purchaser can avoid the supplier receiving ex ante rent and the supplier may choose not to provide either procedure for those for whom net benefit is negative. The optimal specification depends on which of these dominates.

The application just considered corresponds to the upper right box in Table 1. Some examples illustrate the remaining boxes. Consider first a condition such as a malignant tumour that may be treated either surgically (by, for example, an operation to remove the tumour) or non-surgically (by, for example, radiotherapy or chemotherapy). The surgical treatment has a high probability of success in less severe cases but beyond a certain level is less likely to be effective than the non-surgical treatment and its cost increases substantially with severity. The non-surgical treatment consists of a standard course with probability of success decreasing with severity and cost increasing less rapidly than with surgical treatment. For very severe conditions, the probability of any treatment being successful is so low that it does not warrant the reduction in the patient's immediate quality of life. In this case, it is very severe conditions that have negative net benefit from treatment, so higher s corresponds to a more severe case, surgical treatment corresponds to treatment l that is more appropriate for less severe (low s) cases, non-surgical treatment corresponds to treatment h that is more appropriate for more severe (high s) cases, $c_h(s) - c_l(s)$ is decreasing, and $c_i(s)$ is increasing for $i = l, h$. That puts this condition into the upper left box in Table 1.

To illustrate the lower half of Table 1, consider a condition with characteristics similar to those of the preceding example except that the choice is now between surgery alone and surgery combined with non-surgical treatment. An example would again be a malignant tumour but where the treatment choices are surgical removal of the primary tumour either alone or in combination with follow-up radiotherapy or chemotherapy. As in the previous case, higher s corresponds to a more severe condition, surgery corresponds to treatment l that is more appropriate for less severe cases, the combined treatment corresponds to treatment h that is more appropriate for more severe cases, and $c_i(s)$ is increasing for $i = l, h$. The difference from the previous case is that, because the cost of the combined treatment is the cost of surgery plus that of the non-surgical treatment, its cost increases more rapidly with severity than that of surgery alone, so $c_h(s) - c_l(s)$ is increasing, instead of decreasing. That puts this condition into the bottom left box in Table 1.

As an example of the remaining box in Table 1, consider a condition that can be treated surgically at a cost increasing only mildly with severity, managed non-surgically at a cost that increases with severity more rapidly than surgery, or not treated with the patient told to come

back if distressing symptoms recur. An example would be gall bladder disorders for which total cholecystectomy may be expected to leave all severities in a similar condition (no gall bladder) but may be inappropriate for cases of only mild or infrequent distress. Then higher s corresponds to a less severe condition, surgical treatment corresponds to treatment l that is more appropriate for more severe (low s) cases, medical management corresponds to treatment h that is more appropriate for less severe (high s) cases, and $c_i(s)$ is decreasing for $i = l, h$. In this case, $c_h(s) - c_l(s)$ is increasing, which puts it into the lower right box in Table 1.

In the two examples just given, as in the application to tests for coronary artery disease, it follows from the results in Table 1 that supplier rents lead to optimal prices that result in provision of more expensive treatments to patients for whom less expensive ones would be efficient. This illustrates an important point: over-provision of higher-cost treatments to patients for whom lower-cost treatments would be efficient may result simply from an optimal pricing strategy on the part of the purchaser.

In general, optimal prices for the two varieties in the model are different which, in the Medicare context, implies that it is optimal to have the two treatments in different DRGs. The only systematic exception to this when both varieties are provided is when they both have the same cost for all types, in which case they must obviously have the same price if the supplier is to be willing to supply both. In other cases, it is possible for the optimal price to be the same when both varieties are provided but that is essentially coincidental in the sense that the optimal cutoff between recipients receiving one variety and those receiving the other just happens to be at a value of s at which the cost of the two varieties is the same. In general that is not the case because the optimal cutoff depends on relative benefits as well as relative costs. It is also possible for the optimal price to be the same if it is optimal to have only the cheaper variety provided to any recipients. The results here provide a guide as to which price needs to be higher.

Sometimes there are more than two potentially appropriate varieties of provision, which may call for more pricing categories. McClellan (1997) gives the example of patients admitted with ventricular arrhythmia, with three different types of medical treatment: medical management (with three corresponding DRGs in 1993), electrophysiologic stimulation (with one DRG), and automatic implantable cardiac defibrillator placement (with two DRGs). Extending the model to handle this creates no conceptual problem; it just multiplies the number of cases to consider. But, in practice, the administrative costs of having additional pricing categories may limit the extent to which it makes sense to add to the number.

6 Other applications

To illustrate application of the model to a profit-maximising purchaser consider the repair of insured autos mentioned in the Introduction. In that case, the supplier is the repairer who inspects the damaged auto and decides what action to take. The repairer can repair the damaged parts, replace them, or claim that the damage is the result of wear and tear or neglect and

so not covered by the insurance (no provision). Repair is typically more appropriate for less extensive damage than replacement so, with s corresponding to the extent of the damage, repair corresponds to provision l and replacement to provision h , with the cost of repair increasing more rapidly with the extent of damage than the cost of replacement. With either repair or replacement, the insurer's obligation is met in full, so b is the same and independent of s . (It makes no difference to the analysis if the customer pays a fixed deductible. An extension might have the benefit affected by whether the customer is sufficiently satisfied to continue to use the insurer, in which case it might depend on s .) No provision costs the supplier nothing but may result in the insurer facing possible litigation that results in lower expected benefit than repair or replacement. Provided the expected benefit from no provision is lower when damage is more severe (the customer, for example, is more likely to litigate when the repair cost is high), this application falls into the upper left box in Table 1.

To illustrate the case in which the purchaser is also the recipient, suppose the purchaser has a task to be repeated a number of times. The task can be carried out in-house at known cost without specialist equipment (no provision) or outsourced to a monopoly supplier who can use either of two pieces of specialist equipment, one with lower fixed but higher marginal cost than the other. The lower fixed cost equipment requires retaining some in-house staff to do pre-processing work, so the benefit to the purchaser depends on the equipment used. The marginal cost to the supplier of outsourcing with either equipment depends on a parameter s unknown to the purchaser, so the purchaser does not know which method of doing the task is efficient for the number of repetitions required. In-house supply (no provision) is efficient when the marginal cost of outsourcing is high, so higher s corresponds to higher marginal cost. If the equipment efficient for outsourcing with s low has marginal cost increasing more rapidly with s than that efficient for outsourcing with s high, the application falls into the upper left box in Table 1. If less rapidly, it falls into the lower left box.

These applications illustrate how the results established here apply to profit-maximising purchasers, not just those concerned with social welfare. Those results show how optimal provision differs from efficient provision in the face of supplier discretion and how prices need to be set to achieve optimal provision.

Appendix A Endogenous effort to reduce cost

Suppose the cost of providing variety i to a recipient of type s is $\hat{c}_i(s, e) > 0$ for $i \in \{l, h\}$, where $e \in [0, \bar{e}]$ with $\bar{e} > 0$ is the effort (measured, without loss of generality, by the monetary value of its disutility) to which the supplier goes to keep down the cost for this recipient. The cost function $\hat{c}_i(s, e)$ is differentiable with respect to both s and e and both decreasing and strictly convex in e , with $\partial \hat{c}_i(s, 0) / \partial e < -1$ and $\partial \hat{c}_i(s, \bar{e}) / \partial e = 0$, for $i \in \{l, h\}$ and all $s \in [\underline{s}, \bar{s}]$. These properties ensure that higher effort reduces cost and that efficient effort is unique and interior to the interval $[0, \bar{e}]$. Suppose also that the supplier's disutility from providing variety i to a recipient of type s is $\hat{c}_i(s, e) + e$. Thus, if fully reimbursed for the

actual cost of variety, the supplier sets $e = 0$ for both varieties. With a fixed price p_i for variety i , the supplier's utility from providing variety i to a recipient of type s is

$$p_i - \hat{c}_i(s, e) - e, \text{ for } i \in \{l, h\}, s \in [\underline{s}, \bar{s}]. \quad (\text{A.1})$$

If variety i is provided to a recipient of type s , the supplier selects effort $e = e_i(s)$ given by

$$-\frac{\partial \hat{c}_i(s, e_i(s))}{\partial e} = 1, \text{ for } i \in \{l, h\}, s \in [\underline{s}, \bar{s}]. \quad (\text{A.2})$$

This, as is well known to be the case with fixed price payment, is the efficient effort. The assumptions made here ensure $e_i(s) \in (0, \bar{e})$. Now define the functions $c_i(s)$ as follows:

$$c_i(s) = \hat{c}_i(s, e_i(s)) + e_i(s), \text{ for } i \in \{l, h\} \text{ and } s \in [\underline{s}, \bar{s}]. \quad (\text{A.3})$$

The function $c_i(s)$ gives the total cost to the supplier, monetary plus non-monetary, of providing variety i to recipient type s conditional on choosing efficient effort $e_i(s)$. By the envelope theorem, $c'_i(s)$ has the same sign as $\partial \hat{c}_i(s, e_i(s)) / \partial s$, for $i \in \{l, h\}$ and $s \in [\underline{s}, \bar{s}]$. With this definition of $c_i(s)$, all the results in the main text continue to apply.

If, with this extension of the model, $\hat{c}_i(s, e) = \tilde{c}_i(s) - e$, it makes no difference whether the purchaser can observe cost when provision is made and condition payment on that because the supplier can choose e to achieve any desired level of cost without affecting the payoff in (A.1).

Appendix B Proofs

Proof of Proposition 1. The first-order condition that must be satisfied by any \tilde{s}^* at which the maximand in (15) is differentiable is

$$[b(\tilde{s}^*) - (1 + \alpha - \delta)c(\tilde{s}^*)]f(\tilde{s}^*) - \delta[c(\tilde{s}^*)f(\tilde{s}^*) + c'(\tilde{s}^*)F(\tilde{s}^*)] = 0.$$

Equation (16) is a straightforward rearrangement of this. Consider other possibilities for the case $\delta > 0$. For any $s > \underline{s}$, $F(s) > 0$. It follows from (2), (3) and continuity that the maximand in (15) is strictly increased by reducing \tilde{s} for any $\tilde{s} \geq s''$. Thus $\tilde{s}^* < s''$. Given this, the only points at which a maximum may occur with the maximand not differentiable are \underline{s} and s^0 . Under Assumption 1, if s^0 exists it is unique. For $\delta = 0$ (the only alternative to $\delta > 0$), the same argument implies $\tilde{s}^* \leq s''$ and thus does not rule out $\tilde{s}^* = s''$, which will be at a boundary if $s'' = \bar{s}$. But, for $\tilde{s}^* = s''$, the left-hand side of (16) is then zero, so (16) still holds with equality. Finally, for $c_l(\underline{s}) < c_h(\underline{s})$, the supplier chooses variety l for s in the neighbourhood of \underline{s} if these receive provision. It follows from (1) that $\tilde{s} = \underline{s}$ cannot be a maximising value of \tilde{s} because $F(\underline{s}) = 0$, so the right-hand derivative with respect to \tilde{s} of the maximand in (15) is strictly positive. ■

Proof of Lemma 1. It follows directly from Assumption 1 and the definition of \hat{s} in (19) that, for given $s_k^*, \hat{s}^* \in [\underline{s}, s_k^*]$ and the supplier's payoff is maximised by providing types $s \in [\underline{s}, \hat{s}^*)$ with variety $j \neq k$ and $s \in (\hat{s}^*, s_k^*]$ with variety k . (The supplier has the same payoff from providing the two varieties to \hat{s}^* , as well as from providing k and not providing to s_k^* , and so is indifferent as to which is done.) By (1) it is always worthwhile setting p_l sufficiently high that some s receive variety l if none receive variety h , so it cannot be optimal to have no provision to any types. Thus $s_k > \underline{s}$. Moreover, from (3), the left-hand side of (23) is zero for $s_k^* = s''$ and strictly negative for $s_k^* > s''$. But since in all cases $\delta \geq 0$, the right-hand side is non-negative, so $s_k^* \leq s''$, and is strictly positive for $s_k^* > \underline{s}$ when $\delta > 0$, in which case $s_k^* < s''$. For $\delta > 0$, these imply an interior solution for s_k^* , so s_k^* must satisfy (23). For $\delta = 0$, it cannot be ruled out that $s_k^* = s''$, which will be at a boundary if also $s'' = \bar{s}$. But, for $s_k = s''$, the left-hand side of (23) is zero, so (23) still holds with equality. If $\hat{s}^* > \underline{s}$, it is interior and hence must satisfy (22). If $\hat{s}^* = \underline{s}$, the derivative of (21) with respect to \hat{s} must be non-positive at a maximum but, since $F(\underline{s}) = 0$, that corresponds to the left-hand side of (22) being non-positive. ■

Proof of Proposition 2. The proof is numbered in parts that correspond to the parts of the Proposition.

1. Suppose not. Then, either no variety is provided or only variety h is, which corresponds to having variety l being j of Lemma 1 and variety h being k with $\hat{s}^* = \underline{s}$. The former is inconsistent with the result of Lemma 1 that $s_k^* > \underline{s}$. With the latter, by (1), the left-hand side of (22) is strictly positive, which is again inconsistent with Lemma 1.
2. Suppose it is optimal to provide only variety l . This results in a value of the maximand in (21) the same as having variety l being j of Lemma 1 and variety h being k with $\hat{s}^* = s_k^* = s_l^* = s_j^*$. Consider raising s_h above s_l^* . From (21), the derivative of the objective function with respect to $s_h = s_k$ evaluated at $s_h = s_l^*$ is then

$$[b_h(s_l^*) - (1 + \alpha - \delta) c_h(s_l^*)] f(s_l^*) - \delta [f(s_l^*) c_h(s_l^*) + F(s_l^*) c_h'(s_l^*)]$$

whose sign is the same as that of

$$\frac{b_h(s_l^*) - (1 + \alpha) c_h(s_l^*)}{c_h'(s_l^*)} - \delta \frac{F(s_l^*)}{f(s_l^*)}$$

since both $f(s_l^*)$ and $c_h'(s_l^*)$ are strictly positive. But, by hypothesis, s_l^* satisfies (23) with $k = l$. Thus the welfare increase is positive if (24) is satisfied for $s = s_l^*$. Note that, from (2) and (3), the numerator on the left-hand side of (24) is strictly greater than that on the right-hand side for $s \in (s', \bar{s}]$. Moreover, $c_h(s) - c_l(s)$ strictly decreasing implies $c_h'(s) < c_l'(s)$, so the denominator on the left-hand side is strictly less than that on the right-hand side. Thus, (24) certainly holds for all $s \in (s', \bar{s}]$.

3. Suppose it is optimal to have both varieties provided. With $c_h(s) - c_l(s)$ strictly decreasing, it follows from the definition of \hat{s} in (19) that $p_l - c_l(s) > p_h - c_h(s)$ for $s \in [\underline{s}, \hat{s})$

and $p_l - c_l(s) < p_h - c_h(s)$ for $s > \hat{s}$, from which (a) follows directly. With that allocation, condition (2), condition (3) and continuity imply that the left-hand side of (22) is non-positive for $\hat{s}^* \geq s'$. But with $c_h(s) - c_l(s)$ strictly decreasing, $c'_l(\hat{s}^*) > c'_h(\hat{s}^*)$ so the right-hand side is non-negative given $\delta \geq 0$, and strictly positive for $\delta > 0$ and $\hat{s}^* > \underline{s}$. Thus $\hat{s}^* \leq s'$ and from Lemma 1 $s_h^* \leq s''$, both with the inequality strict for $\delta > 0$, as claimed in (b). (c) It follows from (19) that, for both varieties to be provided, $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. With $c'_l(s) > c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(\underline{s}) > c_h(\underline{s})$ implies $c_l(\hat{s}^*) > c_h(\hat{s}^*)$ and thus $p_l^* > p_h^*$. Moreover, it has already been shown that $\hat{s}^* \leq s'$ (with the inequality strict for $\delta > 0$) so, again with $c'_l(s) > c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(s') < c_h(s')$ (or $c_l(s') \leq c_h(s')$ in the case $\delta > 0$) implies $c_l(\hat{s}^*) < c_h(\hat{s}^*)$ and thus $p_l^* < p_h^*$.

4. For $\delta = 0$, (22) is satisfied for $\hat{s}^* = s'$ and (23) for $s_k^* = s_h^* = s''$, for which the payoff W in (20) corresponds to the efficient level in (12) and must, therefore, be globally optimal.

■

Proof of Proposition 3. With $c_h(s) - c_l(s)$ strictly increasing, it follows from the definition of \hat{s} in (19) that $p_h - c_h(s) > p_l - c_l(s)$ for $s \in [\underline{s}, \hat{s})$ and $p_h - c_h(s) < p_l - c_l(s)$ for $s > \hat{s}$, implying the allocation in Part 3(a). With that allocation, condition (1), condition (2) and continuity imply that the left-hand side of (22) is non-positive for $\hat{s}^* \leq s'$. But, with $c_h(s) - c_l(s)$ strictly increasing, $c'_h(s) > c'_l(s)$ so the right-hand side is strictly positive for $\delta > 0$ and $\hat{s}^* > \underline{s}$. Thus, for any interior solution with $\hat{s}^* > \underline{s}$ when $\delta > 0$, $\hat{s}^* > s'$, as claimed in Part 1.

For $\delta = 0$, it follows from (1)-(3) that the only interior solution to (22) has $\hat{s}^* = s'$. But, with $j = h$ and $k = l$, that would imply all types s (except s') receiving the variety that is less efficient for their type. Since the supplier receives no rent in this case, the purchaser's payoff would be increased by having only one variety i provided so that at least some types receive the variety that is efficient for them, completing the proof of Part 1 for $\delta = 0$ and, since $s_i^* \leq s''$ from Lemma 1, Part 2.

It remains to prove Part 3(b). From Part 2, this case can apply only when $\delta > 0$. It follows from (19) that, for both varieties to be supplied, $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. For $\delta > 0$, it was shown in Lemma 1 that $\hat{s}^* \leq s_k^*$ and $s_k^* < s''$, so $\hat{s}^* < s''$. With $c'_l(s) < c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(s'') \geq c_h(s'')$ implies $c_l(\hat{s}^*) > c_h(\hat{s}^*)$ and thus $p_l^* > p_h^*$. For $\hat{s}^* > s'$, again with $c'_l(s) < c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(s') \leq c_h(s')$ implies $c_l(\hat{s}^*) < c_h(\hat{s}^*)$ and thus $p_l^* < p_h^*$. The other possibility for having both varieties supplied in the case $c_l(s') \leq c_h(s')$ is $\hat{s}^* = \underline{s}$ with variety h supplied to \underline{s} only. But this can never be optimal with $c_l(\underline{s}) \geq c_h(\underline{s})$, and thus $p_l^* \geq p_h^*$, because (1) then implies $b_l(\underline{s}) > b_h(\underline{s})$, so it would increase the purchaser's objective to reduce p_h and have type \underline{s} supplied with variety l . ■

Proof of Proposition 4. Part 1. With $c_h(s) - c_l(s)$ strictly decreasing, it follows from the definition of \hat{s} in (26) that $p_l - c_l(s) > p_h - c_h(s)$ for $s < \hat{s}$ and $p_l - c_l(s) < p_h - c_h(s)$

for $\in (\hat{s}, \bar{s}]$, from which (a) follows directly. With that allocation, condition (1), condition (2) and continuity imply that the left-hand side of (29) is non-negative for $\hat{s}^* \leq s'$. But with $c_h(s) - c_l(s)$ strictly decreasing, $c'_l(\hat{s}^*) > c'_h(\hat{s}^*)$ so the right-hand side is strictly negative for $\hat{s}^* < \bar{s}$ when $\delta > 0$. Thus, for any interior solution with $\hat{s}^* < \bar{s}$ when $\delta > 0$, it must be that $\hat{s}^* > s'$. For $\delta = 0$, the unique solution to (29) has $\hat{s}^* = s'$. Moreover, it follows from (26) that, for both varieties to be supplied, $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. With $c'_l(s) > c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(\bar{s}) < c_h(\bar{s})$ implies $c_l(\hat{s}^*) < c_h(\hat{s}^*)$ and thus $p_l^* < p_h^*$. Moreover, with $\hat{s}^* \geq s'$ (with the inequality strict for $\delta > 0$) and $c'_l(s) > c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(s') > c_h(s')$ (or $c_l(s') \geq c_h(s')$ in the case $\delta > 0$) implies $c_l(\hat{s}^*) > c_h(\hat{s}^*)$ and thus $p_l^* > p_h^*$.

Part 2. For $\delta = 0$, it is not optimal to provide both varieties for the same reasons as in the proof of Proposition 3. For $\delta > 0$ and $c_h(s) - c_l(s)$ strictly increasing, it follows from the definition of \hat{s} in (26) that $p_h - c_h(s) > p_l - c_l(s)$ for $s < \hat{s}$ and $p_h - c_h(s) < p_l - c_l(s)$ for $s \in (\hat{s}, \bar{s}]$, implying the allocation in (a). With that allocation, condition (2), condition (3) and continuity imply that the left-hand side of (29) is non-negative for $\hat{s}^* \geq s'$. But, with $c_h(s) - c_l(s)$ strictly increasing, $c'_h(s) > c'_l(s)$ so the right-hand side is strictly negative for $\hat{s}^* < \bar{s}$ when $\delta > 0$. Thus, for any interior solution with $\hat{s}^* < \bar{s}$ when $\delta > 0$, it must be that $\hat{s}^* < s'$. Finally, it again follows from (26) that, for both varieties to be supplied, $p_l^* - p_h^* = c_l(\hat{s}^*) - c_h(\hat{s}^*)$. With $c'_l(s) < c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(\underline{s}) \leq c_h(\underline{s})$ implies $c_l(\hat{s}^*) < c_h(\hat{s}^*)$ and thus $p_l^* < p_h^*$. For $\hat{s}^* < s'$, again with $c'_l(s) < c'_h(s)$ for all $s \in [\underline{s}, \bar{s}]$, $c_l(s') \geq c_h(s')$ implies $c_l(\hat{s}^*) > c_h(\hat{s}^*)$ and thus $p_l^* > p_h^*$. ■

References

- Armstrong, M. and Sappington, D. (2005), Recent developments in the theory of regulation, in M. Armstrong and R. H. Porter, eds, 'Handbook of Industrial Organization', Vol. 3, Elsevier Science, Amsterdam. Forthcoming.
- Baron, D. P. and Myerson, R. B. (1982), 'Regulating a monopolist with unknown costs', *Econometrica* **50**(4), 911–930.
- Chalkley, M. and Malcomson, J. M. (2000), Government purchasing of health services, in A. J. Culyer and J. P. Newhouse, eds, 'Handbook of Health Economics', Vol. 1A, Elsevier Science, Amsterdam, chapter 15, pp. 847–890.
- Chalkley, M. and Malcomson, J. M. (2002), 'Cost sharing in health service provision: An empirical assessment of cost savings', *Journal of Public Economics* **84**(2), 219–249.
- Ellis, R. P. and McGuire, T. G. (1993), 'Supply-side and demand-side cost sharing in health care', *Journal of Economic Perspectives* **7**(4), 135–151.
- Guesnerie, R. (1995), *A Contribution to the Pure Theory of Taxation*, Cambridge University Press, Cambridge, UK.

- Laffont, J.-J. and Tirole, J. (1993), *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge, MA.
- Malcomson, J. M. (2004), 'Health service gatekeepers', *RAND Journal of Economics* **35**(2), 401–421.
- McClellan, M. (1996a), Medicare reimbursement and hospital cost growth, in D. A. Wise, ed., 'Advances in the Economics of Aging', Chicago University Press, Chicago, IL, chapter 5, pp. 149–190.
- McClellan, M. (1996b), Reimbursement incentives and health care costs. Stanford University, Department of Economics.
- McClellan, M. (1997), 'Hospital reimbursement incentives: An empirical analysis', *Journal of Economics and Management Strategy* **6**(1), 91–128.
- McGuire, T. G. (2000), Physician agency, in A. J. Culyer and J. P. Newhouse, eds, 'Handbook of Health Economics', Vol. 1A, Elsevier Science, Amsterdam, chapter 9, pp. 461–536.
- Miller, N. H. (2004), Market structure, commitment, and treatment incentives in health care, Faculty Research Working Papers Series RWP04-008, Harvard University, John F. Kennedy School of Government.
- Newhouse, J. P. (1970), 'Toward a theory of nonprofit institutions: An economic model of a hospital', *American Economic Review* **60**(1), 64–74.
- Newhouse, J. P. (1996), 'Reimbursing health plans and health providers: Selection versus efficiency in production', *Journal of Economic Literature* **34**(3), 1236–1263.
- Patterson, R. E., Eisner, R. L. and Horowitz, S. F. (1995), 'Comparison of cost-effectiveness and utility of exercise ECG, single photon emission computed tomography, positron emission tomography, and coronary angiography for diagnosis of coronary artery disease', *Circulation* **91**, 54–65.
- Siciliani, L. (2003), Selection of treatment under prospective payment systems in the hospital sector. University of York, Department of Economics.