Contracting-out health care services: a conceptual framework

Aidan R. Vining a,*, Steven Globerman b

a Faculty of Business Administration, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
b Department of Economics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Accepted 10 October 1998

Abstract

Contracting-out has become increasingly prominent in the health-care sector. It has been used in activities ranging from ‘internal market’ arrangements in which providers compete for funding from government payers to purchases of medical and non-medical inputs by service providers. While contracting-out arrangements for non-medical services have been widely adopted with apparent success, the contracting-out of medical services has met with criticism. Specifically, prominent ‘market failures’ have been identified which allegedly make contracting-out inefficient and even potentially disruptive to health care delivery. This paper presents and discusses a systematic framework for policymakers to identify and assess potential problems in contracting-out health care services, as well as some generic approaches to mitigating these potential problems. A key to the framework is the notion that conditions contributing to potential market failure problems can often be mitigated by policymakers, and that the strategic choices of policymakers in the ‘first stage’ of the contracting process should include an analysis of how the contracting-out environment can be changed to mitigate potential market failure problems. © 1999 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Health care services; Contracting; Framework

* Corresponding author. Tel.: +1 604 291 5249; fax: +1 604 291 5122; e-mail: aidan_vining@sfu.ca

An earlier draft of this paper was presented at the Conference on Contracting, Centre for Health Economics and Policy Analysis, McMaster University, May 21–23, 1997.

0168-8510/99/$ - see front matter © 1999 Elsevier Science Ireland Ltd. All rights reserved.

PII S0168-8510(98)00056-6
1. Introduction

Health care contracting-out is now fairly common in a number of countries, especially the US, Finland, UK, the Netherlands, New Zealand, Sweden and Canada (e.g. see [1–4]). For example, in 1993 the largest hospital in Canada (the Toronto Hospital) began contracting out a broad range of services including nutrition services, housekeeping, plant operations and maintenance, transportation (both patients and goods), materials management and logistics and laboratory service [5]. Health care contracting-out continues to grow in the US, especially by hospitals, nursing homes and health maintenance organizations. One recent survey found that contracts grew by approximately 20% between 1994 and 1995 [6]. The survey results also show that contracting-out in the US is moving beyond traditional areas such as food preparation, laundry and housekeeping and into clinical areas such as emergency care, substance abuse and skilled nursing, to equipment maintenance and to various management activities (see Table 1). In other countries such as Finland, Sweden and the UK, the major growth in contracting has come

Table 1
Top 20 hospital department contracting out areas

<table>
<thead>
<tr>
<th></th>
<th>Number of hospital clients</th>
<th>Percent change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1995</td>
<td>1994</td>
</tr>
<tr>
<td>Food service</td>
<td>1733</td>
<td>1550</td>
</tr>
<tr>
<td>Emergency</td>
<td>1298</td>
<td>981</td>
</tr>
<tr>
<td>Housekeeping</td>
<td>718</td>
<td>619</td>
</tr>
<tr>
<td>Laundry</td>
<td>557</td>
<td>533</td>
</tr>
<tr>
<td>Clinical/diagnostic equipment maintenance</td>
<td>445</td>
<td>289</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>436</td>
<td>370</td>
</tr>
<tr>
<td>Plant operations</td>
<td>341</td>
<td>310</td>
</tr>
<tr>
<td>Rehabilitation/physical therapy</td>
<td>308</td>
<td>255</td>
</tr>
<tr>
<td>Financial management</td>
<td>474</td>
<td>342</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>211</td>
<td>214</td>
</tr>
<tr>
<td>Skilled nursing/subacute care</td>
<td>134</td>
<td>60</td>
</tr>
<tr>
<td>Security</td>
<td>119</td>
<td>100</td>
</tr>
<tr>
<td>Radiology</td>
<td>69</td>
<td>43</td>
</tr>
<tr>
<td>Gift shops</td>
<td>67</td>
<td>44</td>
</tr>
<tr>
<td>Managed care</td>
<td>47</td>
<td>−</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>47</td>
<td>23</td>
</tr>
<tr>
<td>Accounts receivable</td>
<td>39</td>
<td>31</td>
</tr>
<tr>
<td>Materials management</td>
<td>33</td>
<td>26</td>
</tr>
<tr>
<td>Surgery</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>Anesthesia</td>
<td>30</td>
<td>21</td>
</tr>
</tbody>
</table>

Source, adapted from Moore, 1996:61.

2 The purchase of diagnostic laboratory services in British Columbia is described in Kilshaw, et al. [91].
from the emergence of purchaser/provider splits [7–9], although these splits have resulted in a wide variety of contracting environments [8,10].

In spite of its growth, health care contracting has many critics [11–15]. The major criticisms appear to be: (1) in many circumstances there is limited competition among suppliers which leaves contractors vulnerable to opportunist behavior, such as price gouging [14,16,17]; (2) most medical procedures are complex; this generates two distinct contracting problems. Firstly, it is difficult for contractors (e.g. hospitals) to purchase medical services because there is no ‘best way’ ex ante to provide the service, although there may be ex post. Secondly, for some services, once contractors have contracted the service they no longer have the expertise or knowledge to evaluate the quality of the service provided—only the contractees have this information [14,16]; (3) contractors risk making contracts with contractees that lock them into the relationship, even given unsatisfactory performance [18]. These potential ‘hold-up’ problems are a concern for contractees also, e.g. they may need to make capital investments to provide services exclusively demanded by the contractor.

Complex contracting-out will always present significant risks to public purchasers, but the risks can be mitigated by identifying the nature of the risks and implementing appropriate contracting provisions. At the moment the lack of such knowledge is probably the most serious barrier to effective contracting out. Practitioners appear to have great difficulty in specifying adequate contract terms in health [16,19], as well as in other policy arenas such as education [20]. In other words, government contracting capacity is low, especially in developing countries [16,17].

The paper proceeds as follows: firstly, we present a comprehensive framework for assessing contracting costs and benefits from a societal perspective; secondly, we identify specific types of costs associated with contracting-out; thirdly, we put forward the major determinants of contracting-out costs: task complexity, contestability and asset specificity; fourth, we present a number of contracting-out situations and suggest appropriate strategies for each and finally we present the conclusions.

2. A social efficiency perspective on contracting-out

For contracting-out to save societal resources, several issues must be confronted. Firstly, the objective of contracting-out must be clear; secondly, a framework must be developed, and thirdly, it must be shown that such a framework can be applied to real (and potentially complex) health contracting-out problems. The objective of public sector health agencies should be to minimize the costs of ‘delivering’ any given quantity and quality of service(s). Costs, in turn, consist of resource expenditures for inputs (production costs) and the costs associated with ‘governing’

\footnote{We do not consider whether the government should be providing a particular service in the first place. See [43] on this broader issue.}
the production of health care. The evidence suggests that there is normally a potential for contracting-out to lower the first set of costs (to be discussed below); however, these savings might be more than offset by increases in governance costs. The major purpose here is to suggest a framework to address two fundamental issues: how can government or quasi-government health agencies assess ex ante the potential governance costs that arise with contracting-out? How, and under what circumstances, can governance costs be reduced? The framework identifies alternative instruments that are more or less effective depending upon the relevant circumstances, and provides a basis for meaningful categorization of them.

It is useful to distinguish between two levels, or types, of contracting in health care: (1) specific organizational contracting-out, such as by hospitals for testing and diagnostic services; (2) purchaser (contractor)/provider (contractee) splits (often at the regional level) that are appearing in many countries, hereafter referred to as institutional contracting-out. A number of jurisdictions, including the UK, New Zealand and Sweden, are responding to increases in health care expenditures by regionalizing health budgets. The general strategy is to create ‘internal markets’ by dividing a jurisdiction into geographic regions, set up regional or sub-regional management structures and then allocate a fixed healthcare budget to that entity (e.g. [8, 21]). These regional entities then contract for services. In the US some state-funded services, such as public mental health services, have been delegated to the counties in an analogous manner [22].

The implications of institutional-level contracting-out are still unclear. Because of interest group political pressure to maintain local health care facilities, few jurisdictions have apparently been able to implement effective purchaser/provider splits as yet. For example, under the Area Health Boards in New Zealand there has been the tendency of Area Health Boards to purchase their own services rather than those of private providers. The Area Health Boards had no incentive to contract with private providers as this could result in under-use and eventual closure of the Boards own hospitals [14]. Efforts at contracting in Sweden have also been hampered by the unwillingness of local purchasers to send patients in their administrative regions to providers outside those regions [23], and to an apparent tendency of county councils to favor the incumbents in awarding contracts [8]. In the early stages of reform in the UK, the government was reluctant to permit widespread hospital closures [24]. Nevertheless, evidence from the UK suggests that, given time, new patterns of resource allocation do emerge under institutional level contracting-out [9, 25]. It is likely that the major problem will switch from being resistance to the practical difficulties of effective contracting governance [16, 17].

The underlying premise of the analysis is that contracting-out should be utilized when society is made better off by doing so. This objective is not the same as minimizing governmental costs or a given health unit’s costs. Public sector expendi-

---

4 For an estimation of these ‘organization costs’, see Masten, Meehan and Snyder [84].
5 This framework was originally discussed in Globerman and Vining [90].
tures are only one component of aggregate social cost. Other costs are the costs borne by private and not-for-profit organizations that provide the contracted-out services (hereafter contractees) and costs to citizens (including consumers of the service and third party citizens who are not directly involved in receiving the service). One reason that adopting such a perspective is consequential is that hospitals and other health care providers are being encouraged to be entrepreneurial in controlling costs and providing high quality service. If cost control at the individual hospital levels results in higher costs for other participants in the health care system (including patients), society will either be worse-off, or at best, no worse-off.

3. The costs of contracting-out

Three types of costs are relevant in the choice between internal production and contracting-out: production costs, bargaining costs, and opportunism costs. Bargaining and opportunism are costs of governance. Bargaining costs include the following kinds of costs: (1) costs arising from negotiating contract details per se; (2) the costs of negotiating changes to the contract in the post-contract stage when unforeseen circumstances arise; (3) the costs of monitoring whether performance is being adhered to by other parties, and (4) the costs of disputes which arise if neither party wishes to utilize pre-agreed resolution mechanisms, especially 'contract breaking' mechanisms. While only the first cost is experienced at the time of contracting (the others are experienced in the post-contractual period), virtually all of these costs can be anticipated and dealt with at the time of contracting.

Bargaining costs arise when both parties are acting with self-interest, but in good faith [26]. The incremental bargaining costs of contracting-out are relevant because an advantage of not contracting-out (i.e. ‘internalizing’ the activity) is that the distribution of costs across the organization do not need to be bargained. Bargaining within organizations—for example over wages, bonuses or internal transfer prices—is also costly [27,28]; thus it is incremental bargaining costs which are relevant [29,30].

‘Opportunism’ is behavior by a party to a transaction designed to change the agreed terms of a transaction to be more in its favor. Opportunism costs arise when at least one party acts self-interestedly, but in bad faith. Opportunism is more likely with contracting-out than within organizations, since who gets the profit is more relevant in dealings between organizations. (This applies whether the separate organizations are for-profit or not). Additionally, employees within organizations have better opportunities to ‘pay back’ (and, therefore, discourage) opportunistic organization members [31]. Opportunism can also occur within organizations [32,33]. Therefore, again it is incremental opportunism costs, which are relevant. Opportunism is more often considered to occur after contracting has taken place, but some behavior prior to contracting have ‘opportunism-like’ characteristics [34].

Although analytically it is possible to make a sharp distinction between bargaining and opportunism costs, in practice, they are difficult to distinguish—it is almost
always in the interest of opportunistic parties to claim that their behavior results from an unexpected change in circumstances (i.e. uncertainty). Frequently, the other party cannot tell whether the claim has genuinely arisen from unforeseen circumstances or not. An inability to distinguish between legitimate bargaining and opportunism itself raises contracting-out costs [35].

Production costs are the opportunity costs of the resources—land, labor and capital—actually used to produce the output. Economic theory suggests that production costs will normally be lowered where there is competition. Both theory and evidence also suggest that in a competitive environment private, profit-oriented firms generate lower costs [36,37]. Production costs are likely to be lower with competitive contracting-out for two reasons.

Firstly, in-house production may entail production at too low a level to be efficient [38,39], that is, the public organization does not use enough of the good itself to be able to produce it at minimum efficient scale. An independent producer, selling to multiple buyers can achieve minimum efficient scale. These production costs should be conceived broadly—the most significant economies of scale might be in intangible factors such as administrative systems, knowledge and learning and access to capital markets. Notice that scale economies are not a sufficient argument for private sector production, only for ‘stand-alone’ production that can utilize economies of scale. In practice, however, it is often difficult to design government organizations that can, for example, span several political jurisdictions to take advantage of economies of scale. On the other hand, not-for-profits may be able to compete on this dimension [40].

Secondly, public provision may fail to achieve the minimum production costs that are technically feasible [41,42]. The standard hypothesis is that this results from the fact that government is normally a monopoly, rather than from intrinsic attributes of government ownership per se ([43], but see [37]). Government monopoly blunts efficiency incentives in two ways: firstly, it eliminates comparative performance benchmarks for customers and, secondly, the service is likely to be paid for through aggregate taxes, thereby obscuring the price and efficiency of a government supplier.

The evidence does not support the conclusion (holding other things constant) that a change from non-profit status to profit-oriented status in hospitals leads to efficiency improvements [44,45]. This evidence supports the idea that competition is the crucial element. Profit-maximizing firms in a competitive market will be forced to price at the lowest possible marginal cost, thus eliminating inefficient practices. The objective function of not-for-profit organizations, including non-profit hospitals and physicians, is complex. A number of different objective functions have been proposed [46–49]; thus, there is some theoretical ambiguity about the impact of competition on hospital and other healthcare provider behavior. However, the evidence suggests that competition in health care results in lower costs and/or prices. In particular, there is evidence that greater intensity of competition in the hospital sector lowers prices [50–52]. This evidence is not, of course, determinative of the benefits of comprehensive contracting-out because many of the activities that are potentially contractible are not intrinsically medical, but relate to management, maintenance, technology and other functions.
There is also considerable (and better) evidence from other activities that contracting-out by government to private suppliers does lower production costs [53–60]. British, Canadian and Australian studies tend to find that production cost savings are in the 20–30% range, especially if competitive bidding is used [56,58,61–64]. A recent summary of the contracting-out experience of the 66 largest cities in the US suggests that the annual cost saving was between 16 and 20%; respondents also estimated that contracting improved service quality by between 24 and 27% [60].

Specifically in the health sector Domberger, Meadowcraft and Thompson [65] found that competitive tendering for health-related domestic services reduced costs by between 18 and 27%. In another study that concentrated on catering, the British National Audit Office found that competitive tendering reduced costs by approximately 10% [66]. Milne and McGee [67] found “that the cost of domestic services fell by 17%...; and that the cost of catering fell by 5% and by 11% for labor costs”.

The evidence suggests that much of the cost reduction comes from productivity improvements, which implies that although nominal wages do not change, there are decreases in effective wage rates [16,56,64,67]. Higher internal costs also result from higher pay rates and civil service restrictions that limit productivity growth [68–70].

However, studies that have examined the relative production costs of internal provision versus contracting-out have not included bargaining and opportunism costs, which a priori might be expected to be higher with contracting-out. Additionally, not all forms of contracting-out can be expected to lower production costs, particularly if cost-plus contracts are used [71,72]. The empirical evidence on cost-plus contracting and on cost-plus regulation confirms this.

In summary, public policy should seek the regime that minimizes the sum of production, bargaining and opportunism costs across government (and other governmental contractors, such as hospitals), contractees and third-party citizens for any given level and ‘quality’ of health care activity. Ideally, one then compares the estimated costs with the costs of internalization, that is, the cost of the organization producing the good itself. An important consequence of this criterion is that governments should ideally be as concerned with the costs that they impose on contractees and/or third party citizens, as with the avoidable costs that contractors impose upon governments or their agents. However, at the organizational level—e.g. the individual hospital—it must be recognized that it is almost impossible to expect even non-profit organizations to act with this degree of altruism without appropriate incentives. Even at the level of institutional contracting, regional contractors may have opportunities to export costs to other regions (for example, by various types of ‘creamskimming’). In these circumstances central governments need to construct contract conditions (i.e. design meta-contracts).

4. The determinants of contracting-out costs

Three major factors are likely to determine the sum of bargaining and opportunism costs: task complexity, contestability, and asset specificity. We discuss each of these in turn.
4.1. Task complexity

Task complexity (the terms ‘product’ or ‘service’ complexity could also be used) describes the degree of difficulty in specifying and monitoring the terms and conditions of a transaction. For example, specifying and measuring the quality of food served by a contractee is relatively easy. Specifying and measuring the quality of complex medical services is relatively difficult [16].

The degree of task complexity largely determines: (1) the uncertainty surrounding the contract (this effects both contracting parties equally) which raises the probability that ‘bounded rationality’ will come into play [73]; (2) the potential for information asymmetry (the probability that one party to the contract will have information that the other party does not have); and (3) the probability that there will be externalities that will affect other organizational or health sector activities. Complex tasks involve uncertainty about the nature and costs of the production process itself. They also face more environmental uncertainty (complex activities are more likely to be affected by unforeseen changes in the task environment) [74–76]. Greater uncertainty raises bargaining costs, both during contract negotiations and post-contract; for example, Masten [77] found that more complex components were considerably more likely to be produced internally in the aerospace industry than to be contracted-out.

Information asymmetry occurs when one party has relevant information that the other party does not. While information asymmetry does not always raise costs, usually it does, especially if a contract involves post-experience goods, that is goods where quality is only revealed considerably after the transaction occurs [78]. This is relevant for many health care outcomes (e.g. long term survivability), but less so for many intermediate goods that are inputs to health care such as food services, cleaning, etc. Table 1 shows that approximately half the list consists of activities that do not require medical (or more broadly health-related) expertise: food services, housekeeping, laundry, equipment maintenance, plant maintenance, financial services, gift shops, accounts receivable management, materials management and security. Thus, for such services, any informational asymmetry problems will not be medical in nature.

High task complexity raises the probability that there will be information asymmetry, because it implies specialized knowledge or assets whose characteristics are only initially known to contractees or other experts. Information asymmetry, thus, raises the probability that a party to the transaction can behave opportunistically. Opportunism arising from information asymmetry can occur either at the contract negotiation stage (typically when there is information asymmetry and low contestability) or at the post-contract stage, but is most likely to be significant post-contract. Either contractor or contractee may generate these costs.

Higher task complexity also increases the potential for production externalities, that is the potential for serious disruption to the rest of the organization if the contracted service is withdrawn or degraded [79].

As noted above, it is commonly claimed that it is difficult to specify and monitor the terms and conditions of secondary and tertiary health care transactions. From
the contractors standpoint, these transactions can be complex, both because there is substantial uncertainty surrounding the transaction, and because contractees often have more information about attributes of the relevant transactions. The associated concern is that it may be very difficult for contractors to ensure that the quality of health care services provided is high and appropriate to the needs of individual patients [14].

Difficulties in specifying and evaluating the quality of care provided to individual patients are not trivial; however, the relevant issue is the intrinsic difficulty that contractors can expect in monitoring deviations between expected and actual quality levels supplied by specific contractees to groups of patients over time. The available evidence suggests that the problem may not be as severe as some critics suggest, often because the existence of alternative providers tempers the incentive of any individual contractee to act opportunistically, e.g. to surreptitiously lower the quality of service [2,8].

Some evidence on this point is provided in Keijser and Kirkman-Liff’s [80] review of the Arizona Health Care Cost Containment System (AHCCCS). The authors point out that the competitive bidding process under the System incorporates quality assurance and utilization management mechanisms at the level of both the contracting plans and the AHCCCS administration to assure high-quality outcomes. The detailed descriptions of quality assurance and utilization management mechanisms supplied by bidders are reviewed by physicians working for the state agency. While the state agency had to develop improved management information systems in order to implement the contracting-out process, the bidding system adopted seemed to be reasonably effective at both cost-containment and quality control.

Early quality of care problems in the Arizona experiment were apparently fairly modest. For example, a survey by the Arizona Family Physicians Association of its members found that three-quarters of those who had received medical care were satisfied [81]. In California, by contrast, complaints about denial of treatment and poor quality problems were widespread. Differences in the experiences of the two state regimes may be related to the fact that in California, unlike Arizona, there was no competitive bidding system to select contractees. The California plans could market directly to potential enrollees. Obviously, problems of opportunism can be more readily anticipated under direct marketing schemes.

The above examples suggest that specifying and monitoring quality and other attributes of health care are not intractable6. Whether the benefits associated with contracting-out justify the costs of monitoring is a separate issue. Undoubtedly, there are economies-of-scale in management information systems. This, in turn, suggests that there are economies in consolidating the number of public sector purchasing administration units, all other things constant; however, such consolidation could leave contractees wary about investing in specific assets (to be defined and discussed below), since contractors would enjoy substantial monopsony power and

6 Some additional support for this assertion is provided by the following observation. Where private diagnostic facilities exist in British Columbia communities, many physicians find that the service for out-patient work provided by the private facilities is superior to that provided by the out-patient services of the hospitals, see Kilshaw, et al. [91] (p. 9).
would be tempted to use it when contractees face significant exit costs\textsuperscript{7}. The overall
consequence of consolidation might therefore be to increase, rather than decrease,
overall costs. An alternative approach might involve groups of contractors sharing
the use of a non-profit and independently managed monitoring board that would
evaluate the performance of contractees on a sampling basis and arbitrate disputes
between the parties.

4.2. Contestability

A contestable market is one where only a few firms are immediately available to
provide any given service, but many firms or non-profit organizations would quickly
become available if the price paid by the governmental organization exceeded the
average cost incurred by contractees. For example, the markets for basic accounting
and payroll services are highly contestable as many firms have the basic capabilities
to supply such services, even if they are not currently doing so. The degree of
contestability may, in some cases, be more important than the number of firms actually
providing the service \cite{82}.

In some circumstances the market for the service in question may be competitive—
there may be a considerable number of organizations in the relevant (usually
geographic) market producing the service, or a very close substitute. However,
governmental entities can often reduce production costs by contracting-out activities
for which there is no private market demand and, therefore, no direct competition
\cite{43}. Such goods may involve extensive scale economies and, as a result, there may
be some degree of local, regional, or even national, natural monopoly (this is also
usually related to the extent of asset specificity, to be discussed next). Under these
conditions, there may be no competition. However, if providers are able to switch
production to the good without sinking large up-front costs, there is contestability.

The degree to which the activity being contracted for is contestable affects
opportunism costs. If the market for the activity is contestable, opportunism is reduced
at the contract stage and, potentially, at the post-contract stage. Low contestability
raises different issues in the contract and post-contract phases. During contract
negotiations, a potential contractee in a market with limited contestability is tempted
to offer services at a price above marginal cost (or average cost in circumstances where
average cost is declining for the demanded good). This higher price can be thought
of as a bargaining cost, because it is part of the contracting-out ‘toll’.

At the post-contract stage low contestability increases the risks of opportunism (and
associated costs) facing the other party and possibly third parties, for two reasons.
Firstly, because a contractee cannot be quickly replaced (temporal specificity). Secondly,
because there is a heightened risk of ‘contract breach externalities. This
risk is especially relevant when the contractee provides services that are related
to a network of some kind. For example, a firm carrying out hospital payroll operations
might threaten to withdraw service, jeopardizing the payment of all payroll

\textsuperscript{7} In Sweden, the creation of central purchasing agencies by counties has apparently been of concern
to local providers such as primary care physicians. See Rehnberg \cite{8}.
paychecks. This could effectively shut down the hospital. Situations where organizations fear breach externalities are often defined as ‘strategic’ systems. However, government or quasi-government agencies do not eliminate these externality problems by carrying out production themselves. Government or hospital employees can also opportunistically hold-up their employer by withdrawing essential services (passive breach) or by picketing and various forms of sabotage (active breach).

As noted above, a number of researchers claim that imperfect competition in the supply of many secondary (or hospital) services attenuates the potential efficiency gains from contracting-out those services. In particular, the presence of small and geographically dispersed populations combined with the existence of economies of scale are suggested as limiting the potential number of local contractees, while high sunk costs act as a barrier to de novo physical entry of new suppliers.

The evidence on contestability suggests that government contractors have intentionally or unintentionally contributed to contestability problems. For example, the evidence from both the UK and New Zealand suggests that contestability is not entirely exogenous, especially at the institutional level [14,67]. Specifically, if potential providers perceive that government contractors are soliciting ‘unreasonably low’ bids and/or are arbitrarily requiring them to rebid at lower-than-originally contracted prices, a competitive market may not emerge. Such behavior could thwart competition even in potentially competitive markets [67,81].

It may also be possible in many cases for contractors to enhance competition by expanding the size of the relevant geographic market through ‘exchange agreements’ among themselves, whereby patients could cross geographic (and administrative) boundaries to acquire health care services, and contractors compensate other contractors who must pay for services to those individuals under the formers’ jurisdictions. A number of practical political obstacles to this type of initiative have been noted in the literature. In particular, specific government purchasers are often loath to fund health care services that are provided outside their administrative boundaries, especially if it threatens the financial viability of provider organizations in their own localities [23]. Moreover, patients appear to be reluctant to travel any significant distance for treatment, even when local hospitals have relatively long waiting lists [83]. In the latter context, contestability problems are not so much the result of sunk cost investments, per se, but of the geographical specificity of the relevant assets, i.e. in this case hospitals and physicians.

Another potential approach to mitigating competition problems is for government contractors to own the (sunk cost) assets associated with service provision, and for providers to own relatively fungible clinical and related assets. As a practical matter, this presumably would oblige purchasers to maintain ownership of buildings and relatively specialized and expensive equipment which, in turn, would be leased to providers who successfully won the right to provide service\(^8\). In this

\(^8\)The actual purchase and maintenance of the relevant capital equipment could presumably be contracted-out. Indeed, examples of this type of outsourcing are cited in the literature. See, for example, Solovy [92]. Ironically, in some jurisdictions such as New Zealand, an uncompromising ownership split between purchase and provision of service is seen as necessary to make the internal market arrangement work efficiently.
way, the need for entrants to make large sunk-cost investments might be mitigated and contestability thereby enhanced.

In urban areas where contestability is usually more extensive, government funders may actually be reducing the workability of competition by encouraging specialization among providers along lines of clinical speciality. By making one provider essentially responsible for all cardiac problems and another responsible for all neurological procedures, for example, government purchasers are reducing their abilities to switch providers in the event of unsatisfactory performance. This observation underscores the need for purchasers to think broadly about the cost consequences of specific policies. In the case of policy-mandated initiatives toward increased provider specialization, anticipated savings in administrative and operating costs, along with possible improvements in quality, associated with economies of specialization, may be achieved at the expense of foregone cost savings associated with internal contracting schemes.

In sum, we do not see either economies-of-scale or the need for sunk-cost investments as the main barrier to contestability. In particular, if either patients or service providers are mobile, small population densities need not prohibit competition. If they are not mobile, the problem is better evaluated as one of geographic asset specificity. Indeed, for the remainder of this paper, we assume that contestability can be achieved in all cases, except when patients must be served in a local region and providers are also locationally immobile.

4.3. Asset specificity

An asset is ‘specific’ if it makes a necessary contribution to the production of a good and has much lower value in alternative uses [34]. There are various kinds of specificity including physical asset specificity, location specificity, human asset specificity, dedicated assets [26] and temporal specificity [84,85]. Whatever the form of asset specificity, the issue is basically the same: contracts which require either party to employ assets (usually capital assets, but in some circumstances human capital assets) that have little or no alternative use, that is, are ‘sunk’, raise the potential for opportunism. The contracting party who commits assets is vulnerable to hold-up [86,87]. No matter what prices are agreed to in the contracting stage, the other party can behave opportunistically by reneging and offering lower prices that only cover incremental costs.

In the health policy literature, concerns about asset specificity are largely implicit. One interpretation of the concerns is that temporal specificity, in turn related to location specificity, is the major problem. Specifically, at the institutional level, contractors/purchasers may find it difficult to interrupt the delivery of health care services when contractees/providers are performing in an unsatisfactory manner. It is often difficult to outsource patients even when there are other health care

---

9 As Pirrong [85] notes, sunk costs can generate opportunism even where no contract has been signed, if a party has committed resources whose values would be reduced if the transaction is not consummated.
providers with excess capacity. This can arise when alternate providers are located at any significant distance from the existing provider. It can also occur if providers, for political reasons, are unwilling to contract with providers outside their region. The resulting inertia provides some scope for providers to act opportunistically, e.g. claim that payments are not compensatory and holdout for renegotiated higher payments.

Where contractees own specific assets, they have to worry about the potential for opportunistic behavior by contractors. For example, MDS’s Toronto laboratory is highly automated and processes over 11 million samples a year [88]. Such automation is likely to greatly increase production economies of scale, thus lowering unit product costs, but it makes the contractee more vulnerable to physical asset specificity hold-up. This is because it is extremely costly to physically relocate a large laboratory to escape the opportunism of local purchasers. On the other hand, new automated facilities tend to have much greater capacity which allows them to service many more hospitals at lower cost. Lower production costs can compensate for higher transportation costs. This increases competition and reduces the potential for hold-up.

Although bargaining and opportunism costs can occur during contracting (period 1) or post-contractually (period 2), it is efficient and feasible to address both these costs at the contracting stage (that is, in period 1). The two parties are in a multi-period game [89]. The contractor ‘player’ can anticipate what the optimal strategy in each period of the game will be for the other player (the contractee) and then by backward induction identify its own optimal strategy in each period. For example, suppose the contractor is playing a game where contestability is high in period 1 and but expected to be low in period 2 and subsequent periods. Contractors, therefore, following the analysis presented above, should be able to predict that a contractee will behave opportunistically or generate bargaining costs in period 2 or some subsequent period. The contractor should, therefore, incorporate this expectation into its period 1 strategy. The optimal result is a period 1 contract that anticipates and addresses all potential opportunism costs and bargaining costs.

The practical lesson of a multi-period game analogy is that a contractor can formulate consistent expectations about future developments and plan accordingly. To do so, contractors must think through the factors influencing opportunism and bargaining costs as well as strategies to mitigate those costs. In principle, contractual promises should be implemented in period 1 such that expected (social) opportunism and bargaining costs are minimized, even though they will only be ‘triggered’ by (contingent) events that occur post-contractually. Thus, it is useful to distinguish between ex ante mechanisms and ex post mechanisms, emphasizing that in the case of the latter it is only the ‘trigger’ that is ex post.

5. Contracting-out situations and possible strategies

We now apply the framework to various combinations of task complexity and asset specificity, again emphasizing that we believe any contestability problems are
ultimately co-extensive with asset specificity problems. Hence, addressing the latter problem obviates the former. We consider possible combinations of these two characteristics with the goal of illustrating the conceptual framework rather than providing a definitive guide.

5.1. Low task complexity and low asset specificity

This combination provides the clearest case for contracting-out. It encompasses most, if not all, non-medical health care-related activities at the organizational level, e.g. laundry, housekeeping and food services. It offers the potential for lower production costs for the service, as well as minimal bargaining and opportunism costs. Low task complexity implies that the government has sufficient knowledge and information to specify contract terms precisely (low bounded rationality and/or information asymmetry). With low asset specificity (and resulting high contestability), inefficient or opportunistic contractees can be quickly replaced.

5.2. Low task complexity and high asset specificity

Given low task complexity, problems associated with high asset specificity almost certainly involve high temporal or locational specificity. There are likely to be few efficiency costs arising from high physical asset specificity if the contractor makes the relevant specific investments, as it is not costly to replace the contractee (given high contestability). There can be a problem, however, if the contractee makes the investment. Once the investment is sunk, a contractee is vulnerable to opportunistic hold-up by the contractor, such as a hospital, which could demand that the contractee deliver the goods at a marginal cost. Given that all potential contractees can deduce this as a possible ex post outcome, they will compensate ex ante, in two possible ways: (1) raising the bid price; (2) utilizing a higher cost production technology that requires less physical asset specificity. Either strategy raises aggregate costs, is inefficient and should be avoided.

These problems can be avoided if the contractor owns the specific asset and rents or leases it to the contractee. However, leasing specific assets is not costless. The contractor is now contracting for two activities—the original contracted-for service and the lease contract [61]. Take the example of contracted halfway houses, where the contractor retains ownership of the buildings. The contractee might fully meet its contract obligations in terms of patient care, but (opportunistically) under-maintain the building. The latter problem is presumably readily solved by including a ‘reasonable usage and maintenance’ clause in the lease agreement.

Temporal asset specificity raises problems at both the organizational and institutional contracting levels; however, this problem is relatively easy to deal with at the organizational level. The first possible case is where the contractee fails to provide contracted performance. The contractors usual insurance against the opportunistic exercise of contract breach is an action in tort. However, this is less desirable than having a contract which mitigates incentives for contractee breach. The contractor can, for example, write a contract that contains provisions that backloads payment (contract completion bonuses) and requires performance bonding.
The second possible case is where the contractor wants to terminate because of unsatisfactory performance, but needs to maintain service until a replacement contractee is in place. The concern here is that in this ‘endgame’ the contractee will act opportunistically. The most obvious way for a contractor to mitigate this risk is demand bonding from a winning bidder plus a contract agreement that specifies quick arbitration of a contractor’s claim for the bond because of unsatisfactory contractee performance. In fact, in the case of medical services where breach externality problems are potentially most serious, professional ethics and concerns about protecting reputational capital should mitigate the relevant risks to contractors.

5.3. High task complexity and low asset specificity

This configuration perhaps best characterizes the supply of a wide range of clinical services provided by physicians and other health care professionals. In fact, government funders and hospital providers currently contract (directly and indirectly) with independent health care professionals. An alternative arrangement would involve the formal contracting-out of the management of these professionals. However, the same issues arise under either arrangement. The main problem is high bargaining costs, owing to honest disagreements surrounding ex ante specifications or ex post performance in relationship to ex ante specifications. In particular, disagreements can arise because ex ante specifications are sometimes costly and difficult to write, and (therefore) parties often have difficulty agreeing after the fact about whether the specifications were satisfied, and if not, whether the contractee acted incompetently or negligently. Opportunism should not be a significant problem, since low specificity implies high contestability, suggesting that switching costs will be low for both parties; opportunistic behavior once identified can be easily countered by terminating the contract.

Over time, the accumulation of practice guidelines should substantially narrow the scope for ‘honest disagreements’ about the appropriateness and quality of services provided.

5.4. High task complexity and high asset specificity

The salient difference between this situation and case 2 above is that reliance upon arbitration or other third-party contract enforcement procedures is more problematic. This flows from the fact that it is more difficult for the third party to identify whether contract breach has occurred. One can imagine the difficulties, for example, associated with vetting the performance of a clinic that provides outpatient psychiatric services. Presumably, comprehensive monitoring of the behavior of the clinic would identify blatant incompetence or shirking. However, the need to do so may obviate the net benefits of contracting-out to the clinic.\footnote{Another problem is that the performance of the clinic may not be fully revealed until such time as patient outcomes can be assessed.}
This type of problem has been discussed in the industrial organization literature. The basic solution suggested is that the contractor provides the contractee with an incentive in the form of ‘economic rent’ which the contractee can expect to earn indefinitely in the absence of an identified contract breach. The potential loss of these rents harmonizes the incentives of the contractee with those of the contractor; i.e. to maintain a mutually satisfactory, long-term relationship.

Table 2 summarizes the relevant issues on each of the four cases described. The table focuses on two issues: the dominant problem(s) to be expected and the general solution. The table identifies different environmental contexts and alternative instruments to modify each context; however, the overall framework emphasizes the following steps:

1. Formulate consistent expectations about the uncertainties surrounding the potential transactions at all stages of contract formulation and implementation.
2. Identify the potential opportunism at different stages of contract formulation and implementation, including the underlying sources: contestability, complexity and/or asset specificity.
3. Identify policies to attenuate the opportunism and assess the consequences of the preferred policies for the overall efficiency of contracting-out versus internal production.
4. Implement the relevant policies prior to the initiation of contracting-out.

### Table 2
A summary of contracting-out states

<table>
<thead>
<tr>
<th>Case</th>
<th>Task complexity</th>
<th>Asset specificity</th>
<th>Dominant problem(s)</th>
<th>Solution(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>Low</td>
<td>Few</td>
<td>Rely primarily on contestability via contract termination (i.e. increase potential suppliers).</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>High</td>
<td>Holdup</td>
<td>For physical assets, contractor owns and leases assets; for temporal specificity, backloaded payments, bonuses and bonding. Use of quick arbitration.</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>Low</td>
<td>Honest disagreements about quality and other performance attributes</td>
<td>Where possible, mutually agreed-upon practice guidelines.</td>
</tr>
<tr>
<td>4</td>
<td>High</td>
<td>High</td>
<td>Opportunism by contractee</td>
<td>Harmonize contractor and contractee incentives through ‘rent-creation’.</td>
</tr>
</tbody>
</table>

6. Conclusions

There is increasing interest in contracting-out in health care. Indeed, there is already considerable contracting-out of non-medical services in the sector. Contracting-out of inherently medical activities has met with substantially more criticism. The main criticisms relate to what economists call non-contestability, task
complexity and asset specificity. Upon analysis, we argue that non-contestability itself reflects asset specificity, particularly specificity relating to the timing and location of service delivery.

In this paper, we suggest that the broad problems associated with contracting-out inherently medical activities can be mitigated by contractual and related strategies on the part of the contractor. We propose a simple framework that relates some alternative strategies to archetypal problem situations surrounding health-care contracting.

Acknowledgements

The authors would like to thank Betty Chung for word-processing assistance.

References