

# Appendix F The economics of social services

## Key points

- Social services have particular features that distinguish them from goods and services that are more commonly traded in markets. Centralised government provision is a traditional response to those features.
- Government's role accentuates the importance of the institutional environment in which social services are delivered.
- Social service delivery, whether by government or non-government organisations, involves multiple and complex *principal-agent relationships*, including multi-tasking by multiple agents answerable to multiple principals.
- Providing agents with good incentives is therefore complicated, and needs to be determined in the light of the specific social service in question.
- Particular issues in the provision of social services include the risk that investments in social services assets are undermined by government decisions taken after those investments have been made (*hold-up*). They also include conflicts between achieving quality, which can be hard to measure, and more measurable activities such as cost reduction (*quality shading*).
- Different organisational forms affect the nature of these risks, as well as how best to address them, and therefore the relative attractiveness of service delivery by such organisational forms.
- Government's dominance in funding, providing and purchasing social services limits the tools available for introducing choice and innovation in social services.
- Government brings certain advantages to social service delivery – relative to for-profit (FP) firms – in terms of *mission orientation*, though to a lesser degree than not-for-profit (NFP) social services providers.
- This means government and NFPs offer relatively high non-monetary incentives to employees, and can commit to relatively low-powered financial incentives, relieving quality shading problems.
- FPs address quality shading issues in other ways, such as through reputation concern, or by adopting team-based production and relative performance evaluation to reduce incentives.
- Hierarchical service provision is useful when the benefits of economies of scale and coordinated decision making outweigh the disadvantages of losing control through delegation and distancing senior decision makers from clients.
- Public purchase typically requires competitive tenders and associated high-powered incentives; for example, due to transparency and accountability requirements.
- However, low-powered incentives can be preferable, particularly when quality is hard to contract. Also, alternative purchase mechanisms such as negotiation or relational contracting offer benefits in certain circumstances.
- Both government and NFPs suffer disadvantages relative to FPs in terms of innovation.

The economics literature may not be an obvious source of insight into improving the effectiveness of social services. For many, *economics* has strong associations with for-profit (FP) firms and with competitive markets for consumer goods, labour and financial capital. Social services appear divorced from such FPs and markets.

However, the economics discipline is concerned with explaining the world as it is, rather than simply exploring abstractions from reality. So the discipline has devoted much thought and attention to the study of social services specifically; as it has to other “markets” that share complex features with social services. Further, many of the issues that arise in commissioning, purchasing and contracting – central topics for this inquiry – are common between social services and other types of services. Those issues are well-studied in the economics literature.

This appendix presents a survey of economic literature relevant to this inquiry. It is necessarily incomplete, as the theoretical and evidential base is vast. For the most part, the appendix concentrates on the issues that arise when contracting between organisations of different types, particularly in the presence of difficult-to-measure objectives. Little of the material is specific to New Zealand or to specific types of social services.

The appendix is also incomplete, as an economics lens is only one possible lens that could be applied. The main text of the inquiry report draws from multiple literatures and multiple case studies, and synthesises a view across multiple lenses. This appendix is presented for readers specifically interested in the economic literature that has contributed to, but not determined, the inquiry’s analysis and recommendations.

## F.1 About this appendix

This appendix uses relevant economic theory and evidence to help understand social service delivery. It provides a framework for assessing the benefits and costs of delivering social services through markets, governments or non-government organisations (NGOs), including not-for-profits (NFPs) and FPs. This framework helps in understanding the delivery of social services as it is, and also in thinking about how to improve existing delivery models.

Social services are varied and complex; so, in general, there will not be just one best way of delivering all such services. Further, how a particular social service is best delivered will change over time in response to both changing needs and changing technologies (eg, in information technology (IT) and contracting).

The appendix emphasises factors affecting the incentives of parties purchasing or delivering social services, as well as those of social services clients. The features of social services that distinguish them from the delivery of other types of services are highlighted, as are the rules – that is, *institutions* – affecting either government or NGOs. Other highlighted issues include whether social service providers should compete or collaborate, the importance of a social service provider’s personal motivation – or sense of mission – in providing its services, the implications of providing social services clients with greater choice, and also how best to encourage innovation in service delivery.

Whether or not state assets should be privately owned – that is, *privatised* – is not the subject of this appendix. However, the ideal ownership of assets involved in delivering social services, such as service providers’ client relationships and intellectual property, is discussed as it relates to ensuring good incentives for service delivery.

The questions of whether governments should *make or buy* goods or services, and how best to purchase them (ie, *how to buy*), are well-treated in the economics literature. The literature also has insights into how these decisions should be exercised in relation to social services in particular. There is recognition in economics of the roles and importance of social services delivery by NFPs as well as FPs and governments. However, understandings are still developing as to how these alternative organisational forms affect the make or buy decision, the best way to purchase services when they are not delivered by government, and both competition and innovation. This emphasises that the “technology” available to understand social service delivery has evolved – meaning that traditional delivery models need no longer apply.

The appendix begins by setting out the distinctive features of social services and the environment in which they are delivered (section F.2). This necessarily involves discussion of the costs and benefits of providing social services directly by governments. It then discusses the key factors affecting whether social services are better delivered by governments or NGOs (whether NFPs or FPs). The section explores the best level and form of incentives to encourage cost-effective, quality and innovative delivery of social services.

Section F.3 addresses related questions, such as who should decide whether governments should *make* social services in-house or *buy* (ie, purchase) them from non-government providers.

How best to purchase social services, if they are to be privately provided, is addressed in section F.4, as are issues of whether providers should compete or collaborate, and whether services should be bundled or separate. Section F.5 discusses encouraging innovation in the delivery of social services. Section F.6 contains short discussions of the impact of increased consumer choice on provider market power, and of collaboration between providers.

## F.2 Distinctive features of social services

### Distinctive features of “social” services

In general terms, social services involve the delivery of goods or services. Examples include providing food from a food bank and assisting unemployed people into work. As such, social services can be analysed in economic terms like other goods or services. However, there are distinctive features of social services that require particular attention, and which affect how they should be analysed and how best they can be delivered. The features discussed below will each arise to varying degrees in any given social service, so any analysis of how best to provide a particular social service will ultimately depend on its particular characteristics.

#### Merit goods

One important distinctive feature is that many social services are regarded by modern societies as a form of *merit good*. This means they are something that people should be able to receive aside from their willingness or ability to pay, and instead should be available on the basis of their need. Equity of access is therefore an important consideration in social service delivery, and motivates interventions in the form of public *funding* and possibly also *provision* of social services.

#### Spill-over effects

Related to this is the fact that many social services create *social* benefits or costs that can differ markedly from their *private* benefits or costs.<sup>1</sup> This means that private individuals might over-consume “social bads”, such as through excessive alcohol consumption, if they consider only their private costs. Conversely, they may choose to under-consume “social goods” such as drink-driver re-education, which benefit family members and other road-users as well as the consumer, if they consider only their private benefits. Indeed, social services clients might even be unwilling or hostile to consuming such services, and instead may need to be coerced to do so (eg, by the courts).

Since the cost of accessing services with strong social benefits may deter some from consuming them even when they are willing, this further motivates public *funding* and possibly also *provision*. Indeed, where even free social service delivery does not induce the socially-desired level of social service consumption, non-price means such as social marketing campaigns or coercion may be necessary.<sup>2</sup>

#### Public goods

Private under-provision can become even more pronounced when the services have *public good* attributes.<sup>3</sup> Public goods have features meaning they are consumed collectively, and include examples such as population immunity to communicable diseases (eg, measles).<sup>4</sup> Because of these features, self-interested private providers are unlikely to have sufficient incentives to provide these goods (here, services) at the socially-desirable level, with market-based price signals under-representing social benefits (Besley & Ghatak, 2003). Public intervention is commonly relied upon to ensure that either these signals are corrected, or that

<sup>1</sup> In other words, they create positive or negative *externalities*. Note that social benefits (costs) also include private benefits (costs).

<sup>2</sup> In principle, government could subsidise social services to the extent that consumers face a negative price – that is, are paid – to consume the relevant services. However, this raises issues of political acceptability, as well as the risk of perverse incentives. For example, paying repeat drink-drivers to attend drink-driving re-education classes might induce such drivers to continue to drink and drive.

<sup>3</sup> Specifically, consumers cannot be excluded from consuming them, and the consumption of a public good by one consumer does not diminish the ability of other consumers to also consume that good.

<sup>4</sup> Other examples include flood control and street lighting.

service provision otherwise achieves the desirable level. However, cooperative non-government approaches are sometimes also feasible (Ostrom, 2000; section F.3).<sup>5</sup>

### Insurance characteristics

Another example of private under-provision relates to social services with insurance characteristics, such as ambulance or fire services. Private insurers might find certain classes of consumer to be unprofitable to insure, and so not serve that market. Alternatively, they might only serve those consumers at such a high price as to exclude those lacking the necessary financial resources. Alternatively, some consumers might opt not to take out insurance even when they can afford to. In such circumstances compulsory insurance, with partial or complete public funding or provision, can be important for ensuring equity of access, and reducing any undesirable social costs from consumers having inadequate insurance (Barr, 2012).

### Specific investments

Social service delivery sometimes requires relatively few large-scale investments in fixed assets with specific uses, such as in industries like electricity generation or car manufacturing. However, some services such as schools, hospitals and nursing homes, require significant investments in fixed assets that may have limited alternative uses. Also, where social services rely on large-scale IT, this can involve large investments in customised computer software, and possibly also hardware. Social services often require investments in specific human capital (ie, worker skills), as well as relationship-specific investments (see Box F.1). The latter includes getting to know the needs of particular social services clients, and earning their trust and confidence (which shares features with other industries, such as insurance sales or financial advice).

### Differentiated providers

Because of these specific investments, social service providers are not all alike, offering differentiated skills and knowledge. This in turn means they enjoy a degree of market power, with the result that they do not intensively compete on cost to deliver standardised services. Providers instead rely on their special features to offer them a competitive advantage.<sup>6</sup>

### Client information

Social services clients can be unclear about the variety and quality of services on offer, or unclear as to which services might best meet their needs. Clients may also be unclear about their needs. These features are shared to varying degrees with other service markets (eg, legal advice). In such markets, consumers face search and switching costs, and might only learn about service quality after using the service, by which time it might be too late (or costly) to change providers (see section F.6).

### Bundling

Another distinctive feature of social services delivery is that it often involves a bundle of related services. For example, a family in need might require food from a foodbank, work re-training and career guidance, support to combat domestic violence or substance abuse, and medical interventions for chronic illnesses. Successfully assisting such clients requires a range of services that can be either mutually reinforcing or in conflict. Conversely, providing some services without other services risks undermining the effectiveness of the services provided. The delivery of social services can therefore require multiple interventions by one or more providers. This raises questions about how best to bundle provision by different providers, and the range of services to be offered by any one provider.

## Distinctive features of the delivery of social services

### Traditional versus contemporary approaches to delivering social services

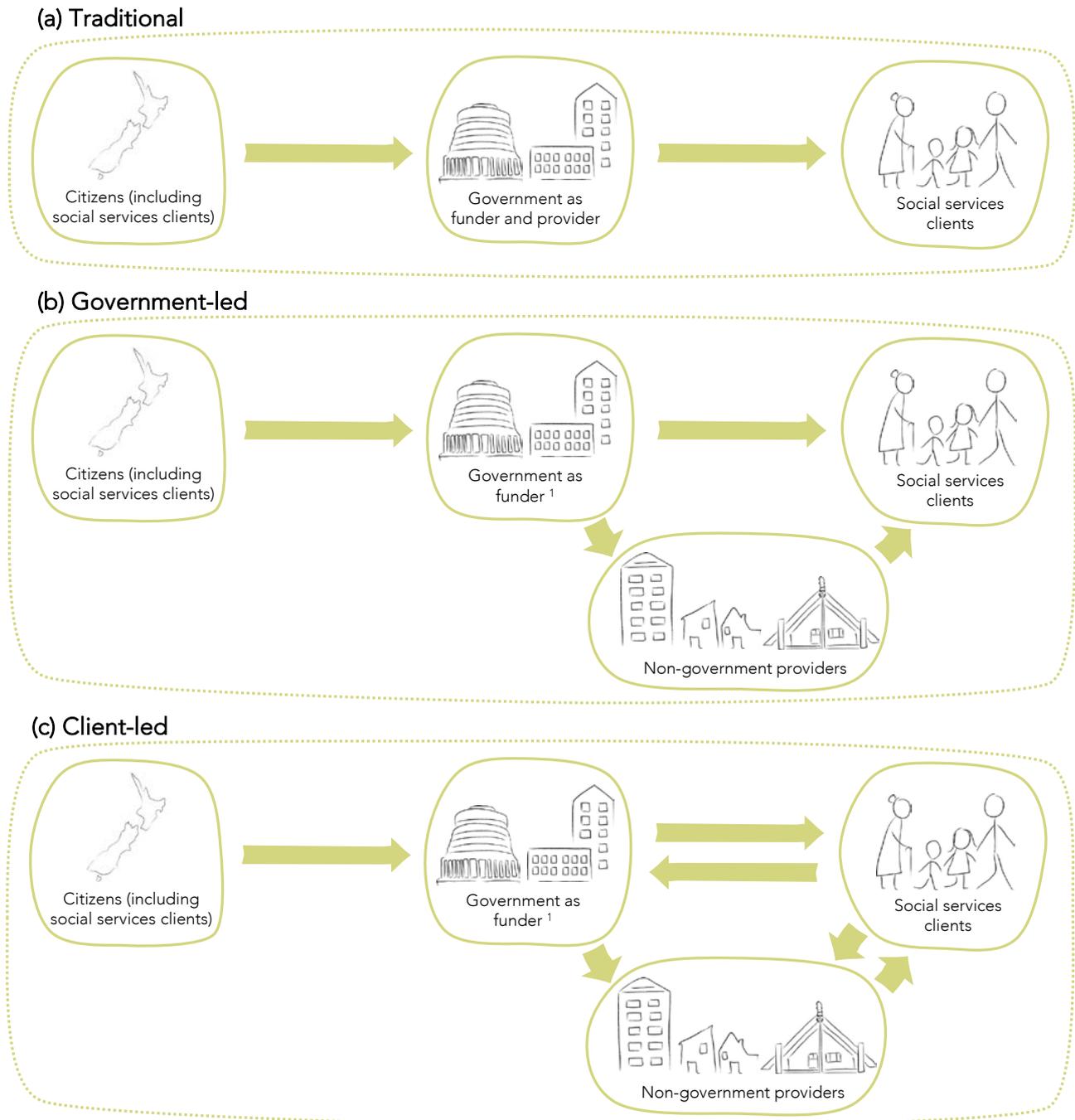
Approaches to organising and delivering social services commonly reflect the above distinctive features. Figure F.1 illustrates three approaches in highly stylised form. The arrows in the Figure indicate who is seeking to induce the actions they desire and from whom.

<sup>5</sup> Ostrom emphasises that market under-provision can also arise for *common pool resources*. Unlike public goods, these resources are *rivalrous*, in that their consumption by one consumer reduces their availability for consumption by others. Social services, such as hospital care, can be better regarded as common pool resources than public goods, because the ability to supply hospital services is limited.

<sup>6</sup> The implications of this market power depend on other contextual factors, such as the extent to which government exerts countervailing market power through its role as primary purchaser of social services, as well as regulations or laws affecting competition.

First is the traditional, government-led approach, in which government acts as both funder and provider of social services. While social services clients as voters exercise broad control over social services, the model typically involves highly centralised and provider-driven decision making (ie, by government), including as to the level, choice and quality of the services provided to social services clients. Little emphasis is given to either client choice, or competition in provision regarding cost, variety or quality.

**Figure F.1 Approaches to delivering social service**



Source: Besley and Ghatak, 2003; Productivity Commission.

Notes:

1. In some cases government might also provide social services, in competition with non-government providers
2. Arrows indicate who is attempting to induce the actions they desire and from whom.

Second is a modernised, yet still government-led, approach. The key difference is that government retains a dominant role in funding social services (and possibly also in providing them), but competing external

providers are now used. These external providers could be NFPs or FPs, with their services usually obtained through either contract tendering or negotiation.

Social services clients continue not to have significant decision rights in this approach. However, depending on how external providers are contracted, greater diversity of services might be available even if clients do not choose their providers. Further, if contracts are tendered, this induces competition on both cost and quality (though, as emphasised later, not necessarily improving quality). However, the competition is directed towards government's preferences, and only indirectly towards those of clients.

Third is a client-led approach, in which social services clients exercise greater control over which providers they use. This causes providers to compete to better match client preferences. However, providers might also compete to meet government's preferences regarding cost, variety and quality, since funding is still provided by government. This represents a hybrid – or *quasi-market* – approach, between public and private provision (Le Grand, 2003). Competition becomes more clearly client-oriented when governments devolve spending decisions to clients: for example, funding *vouchers* of some variety that can only be applied to purchase approved social services.<sup>7</sup>

### **Multiple and complex principal–agent relationships, as well as multi-tasking**

All three models emphasise how decision rights are held by multiple parties at multiple levels. There are parties – who can be called *principals* – whose interests are to be served. There are also parties – who can be called *agents* – acting on behalf of those principals, but who might also wish to act in their own interests instead of their principal's interests. Sometimes a party might be a principal in one role, and an agent in another.

For example, citizens (ie, wider society, including social services clients who can vote) are principals in their relationship with government, which they elect to act in their interests. However, government as funder and/or provider of social services in turn acts as principal in relation to those clients. This arises, for example, when public accountability for taxpayer funds, or responsibility for delivering society's wishes, requires government to induce social services clients to act in certain ways (eg, participate in drink-driver re-education). This is particularly the case where those services are imposed on clients (eg, by the courts), or as a condition of receiving income support.

Parties can be an agent of more than one principal. For example, in the third, client-led approach, service providers are to serve the interests of both social services clients and government, and so are agents of each such principal – a situation known as *common agency*. This “servant of more than one master” feature of the delivery of social services generally weakens incentives (section F.3).

The picture is further complicated by the fact that government comprises multiple, hierarchical, principal–agent layers. These include accountability relationships between the Prime Minister and Cabinet, between Cabinet ministers and ministerial heads, and (below those heads) between managers and line workers. These “tiered principal–agent relationships” also give rise to specific incentive issues, such as how to avoid collusion among lower-level agents, or even encourage it when beneficial (section F.3).

Government is also organised into multiple, parallel ministries, more than one of which might ultimately deliver social services to a particular consumer. For example, one household might receive education, health, job training and/or law-enforcement services. Further, any one of such social services interventions potentially involves multiple tasks – that is, *multi-tasking*. For example, health-related social services might include sickness prevention as well as rehabilitation for past accidents.

When (multiple) agents are involved in multiple tasks, this makes it difficult to create proper incentives for those agents. This is especially the case when the outputs of some of those tasks cannot be easily observed (section F.3). For example, strong cost-reduction incentives might conflict with the quality levels that principals desire if quality is hard to measure. This would favour weakened incentives for cost reductions.

---

<sup>7</sup> The arrangements can be explicit or implicit. Implicit voucher funding arises, for example, where social services clients are free to choose their provider, with government funding then following that choice of provider. In this case, no explicit funding vouchers are allocated.

Interactions among tasks also affect whether they should be provided together or apart, or competitively or collaboratively.

Figure F.1 highlights three main types of party involved in delivering social services – funders, providers and direct clients. To the extent that social services include significant social costs or benefits over and above those enjoyed by direct social services clients, a fourth party type could be added. In particular, this might be fourth-party beneficiaries (in the case of additional social benefits) or fourth-party sufferers (in the case of additional social costs). Those parties can be concentrated or diffuse, as well as identifiable or anonymous. For example, victims of domestic abuse can be concentrated and identifiable, while the beneficiaries of education are diffuse but identifiable.

### **Government as a dominant purchaser and/or provider**

Implicit in Figure F.1 is that government is a dominant purchaser and/or provider of social services. In the traditional approach it is in fact the only purchaser and provider – effectively a state monopoly. Price signals that would normally be available to influence consumer choices in markets are replaced by bureaucratic choices affecting the cost, variety, quality and quantity of services consumed.

In all three approaches government retains a prominent role in funding, either compensating for a lack of – or crowding out – coordinated private funding. In the second and third approaches, government might compete with non-government providers to some degree. However, due to its dominant role as purchaser of those services, government dictates the terms of that competition, administratively specifying the quantity, variety, cost or quality of non-government providers' services.<sup>8</sup> As such, simplified "textbook" economic models for analysing the effects of competition require adaptation for the social services environment, reflecting that the social services "market", in this case, is largely driven by administrative choices made by government.<sup>9</sup>

### **Distinctive features of the social services delivery environment**

From the discussion above it is clear that the delivery of social services differs from the delivery of services in a more market-based environment. In that environment, consumer choice and competition (on price, variety and quality) play key roles. Conversely, in social services the roles of government as funder, provider and/or buyer of social services (from external providers) are prominent. This leads to a discussion of how the political and wider institutional environment in which governments operate affects the delivery of social services.

#### **Political environment**

Three factors stand out as particularly important in affecting government's role as funder, provider and/or buyer of social services. The first is the electoral cycle, and within that the ministerial cycle. The second relates to opportunistic behaviour, either by governments or third parties. The third is the "machinery of government", including the nature of Cabinet and budget rules.

#### **Electoral cycle**

The electoral cycle places a natural constraint on the horizon over which politicians take decisions regarding the objectives, resourcing and delivery of social services programmes. In New Zealand a government is potentially at risk of losing power every three years, although it might expect to survive more than just one electoral term. This naturally causes politicians to be wary of programmes with large short-term costs and/or uncertain long-term gains, even if such programmes could be justified in terms of improving social outcomes.

Programme delivery can suffer from short-termism as a consequence, unless governments have the means or commitment to implement long-term programmes. Such means include being able to secure bipartisan

<sup>8</sup> Where private providers are involved in the delivery of social services, government cannot dictate all relevant terms of such private provision. For example, if it calls for tenders for a specified quantity, variety and quality of social services, private parties remain free to choose the cost at which they are willing to provide those services.

<sup>9</sup> These features arise to greater or lesser extents in other specific markets. There is much collective research and expertise in extending textbook models to incorporate such peculiarities.

support for long-term programmes such as retirement savings.<sup>10</sup> They also include governments having a very high chance of re-election, meaning they can plan on programmes spanning more than one electoral cycle. Legislation embedding long-term decision horizons, such as New Zealand's Fiscal Responsibility Act 1994, also aids in implementing long-term programmes. A commitment to long-term programmes can also arise to ensure delivery on flagship election promises, or simply as an attempt to tie the hands of successor governments.

Indeed, the electoral cycle raises the risk of a change in government, with an attendant risk of change in social policy. A long-term social service programme favoured by one government might be interrupted mid-delivery with the arrival of another – an example of hold-up risk (see Box F.1). This fundamentally affects the incentives of parties – government or non-government – to make the long-term or specific investments required for delivering social services (eg, in human capital, service models and infrastructure, community knowledge, and client relationships).<sup>11</sup>

#### Box F.1 **Hold-up**

*Hold-up* refers to a risk confronting parties that invest in assets which have value in a particular use or in the context of a particular relationship, but a much reduced (or even zero or negative) value otherwise. Such assets are called *specific assets*. Examples include tooling up machinery or hiring specialised staff to supply goods or services to a customer, and learning about the specific needs and circumstances of customers.

Once these investments are made, the customer can seek to renegotiate the contract terms. Renegotiation (or the potential for it) undermines the value of the investment to the party making it. Anticipating this risk, the investing party may under-invest relative to the efficient investment level. One solution is for the parties to *vertically integrate* – that is, for the party at risk of being held up to own the other.

Hold-up is relevant to the provision of social services if a party supplying social services makes specialised, especially long-term investments (eg, in staff training and IT infrastructure) whose value can be undermined by decisions taken by the party purchasing the social services after those investments have been committed to.

*Source:* Williamson, 1999; Grossman and Hart, 1986.

### Ministerial cycle

As well as the electoral cycle, there is also a ministerial cycle. This refers to the expected tenure of a government minister, which might span multiple electoral cycles, but which might be shorter than a single electoral cycle. The latter is the case for poorly performing ministers, but can also be the case when unpopular governments (or prime ministers) seek to present a “fresh face” to the electorate.

Indeed, changes of either governments or ministers often lead to programmes being replaced or rebranded, regardless of whether existing programmes are fully implemented or successful. In either case, this exacerbates hold-up risks for long-term programmes involving specific investments. However, many government programmes in fact survive ministerial (or even governmental) changes. So these risks are most relevant to programmes that are particularly linked with ministerial (or governmental) tenure.

### Opportunism by governments or third parties

The second factor, opportunistic behaviour, also presents service providers with hold-up risks. The first type – *government opportunism* – refers to the fact that, absent other safeguards, governments are relatively free

<sup>10</sup> In stable democracies such as New Zealand, there can be a broad political consensus on major issues. Such issues might be whether or not to provide social services, and the role of markets in the economy. However, political differences can persist regarding finer details, such as the relative size of particular programmes, the extent of market regulation, and so on.

<sup>11</sup> By *specific investments* it is meant that the assets created by the investment have value in their specific use (eg, in the context of a specific relationship either with government, consumers or other providers), but little realisable value otherwise.

to change earlier decisions made regarding the delivery of social services.<sup>12</sup> They can even pass new laws or regulations to do so. Such changes might be insignificant, or even beneficial in light of new information. However, they also have the potential to seriously undermine previous long-term or specific investments in programmes that deliver social services, when those investments were predicated on those earlier decisions.<sup>13</sup> When parties making social services investments anticipate such hold-up risks, they invest at less than the socially optimal level.

The second type – *third-party opportunism* – refers to the fact that interest groups, political opposition parties and rival social services providers have incentives to highlight – or create the perception of – flaws in delivery programmes (Spiller, 2008). This might be on philosophical grounds, to secure changes that better reflect their own preferences, or to embarrass the Government. Since this creates electoral risk for governments, third-party opportunism can induce governments to change social services programmes mid-stream, creating hold-up risks.

Additionally, third parties might challenge the payment of performance-based rewards to service providers, even when the criteria for such rewards have been satisfied. This too creates hold-up risks for providers. Both governments and providers can find it preferable to minimise the risks of third-party opportunism, when anticipated, by distorting social services delivery programmes. They do this by increasing the prescriptiveness of social service contracts, and the rigidity with which those contracts are implemented. In turn, this contractual prescriptiveness and rigidity distorts social services investments.

### **Machinery of government – Cabinet and budget rules**

The third type – *machinery of government* – also affects the delivery of social services. This includes the nature of Cabinet, as well as the operation of budget rules.

Cabinet in New Zealand comprises the subset of government members of Parliament who have been allocated responsibilities for managing different areas of government activity.<sup>14</sup> These ministers meet under the direction of the Prime Minister, and operate under both formal and informal Cabinet rules, as well as relevant laws. Cabinet members may be assigned one or more portfolios of responsibilities, and either principal or associate-level responsibilities for each. Ministers are accountable for their discharge of those responsibilities – to the Prime Minister and Cabinet, to Parliament, and to the electorate.

Cabinet is the forum where decisions are taken regarding government activities, coordinating at a high level across competing demands for limited budget funds, with the Prime Minister and Minister of Finance playing key roles.<sup>15</sup> It is therefore important for deciding how the activities of different government agencies are to be bundled or unbundled. This involves trade-offs between achieving desirable mixes of social services, retaining clear lines of accountability (between government officials and ministers, and between ministers and Cabinet), and ensuring that the provision of each activity is organised efficiently.

The competition for funds at Cabinet level has the potential to improve the information available for choosing funding priorities, which is discussed further in section F.3. Likewise, how different activities are grouped together can either give rise to advantages or disadvantages in creating good performance incentives.

Budget rules play an important role in shaping how Cabinet-level trade-offs are made. To ensure transparency and accountability in how taxpayer funds are spent – particularly given the risk of third-party opportunism – government activities and responsibilities are grouped into defined task areas that are easier to monitor and control. These *silos* need not align with the bundles of activities that matter to the clients of social services. This creates costs for those clients in terms of having to deal with multiple silos, and also in terms of poorly coordinated service delivery. This means that government involvement in those services

<sup>12</sup> See the following subsection on institutional environment for a discussion of safeguards.

<sup>13</sup> This can remain the case even when investments were contracted for, due to the cost, risks and uncertainties associated with contractual enforcement (eg, through the courts). It is particularly so when governments are able and prepared simply to legislate changes affecting previous investments.

<sup>14</sup> Some ministers may sit outside Cabinet.

<sup>15</sup> The Prime Minister has a degree of autonomy to take decisions outside Cabinet, as has the Minister of Finance in respect of public finances.

trades off service coordination against the need for transparency and accountability in the use of taxpayer funds in a political environment.

Further, the need for transparency and accountability under budget rules means that attention of both ministers and third parties focuses on the dimensions of government activity that can be readily measured. One consequence of this is the possible neglect of social services dimensions that are not easily measured. Indeed, in many government activities there are no objective measures of outcomes, and even inputs into those activities can be hard to measure. As a consequence, social services programmes are often assessed in terms of budgetary commitments (ie, dollars spent) rather than in terms of social services outcomes.

Related to this is that where performance incentives are implemented, they are naturally based on measurable dimensions since they cannot be based on dimensions that cannot be measured. Strong incentives on measurable dimensions cause priority to be given to the delivery of those dimensions. But this can cause the poorly measured dimensions to be neglected or undermined, which is emphasised in the discussion of multi-tasking in section F.3. And when delivery is measured simply in terms of dollars spent, this may bear little relation to actual outcomes achieved.

### **Institutional environment**

Related to the above discussion of political environment, the wider institutional environment has the capacity to improve or undermine the delivery of social services. Here *institutions* are taken to mean the “rules of the game”, constraining and shaping the conduct of both governments and non-government parties (North, 1990). They include formal and enforceable rules such as constitutions, laws and regulations, as well as informal rules based around social and cultural norms of behaviour (including codes of conduct). Notably, in the New Zealand context they also include the Treaty of Waitangi.

#### **Legal institutions**

Institutions critically constrain the extent to which the political environment affects the delivery of social services. In particular, constitutions can protect the value of third-party investments, requiring governments to compensate those making such investments if their value is affected by policy changes. This provides a constraint on policy changes. Constitutions can also entrench social services features such as rights to services and equity of access, which further constrain policy changes.

Strong contract laws and an independent judiciary to enforce contracts in the event of breach also provide constraints. This is particularly relevant where the delivery of social services is purchased by government from non-government parties via service contracts. Relatedly, laws requiring fiscal transparency and accountability of governments can shape the ways in which non-government parties provide social services, such as by insisting on the use of transparent competitive tenders rather than private negotiation (which can be vulnerable to favouritism) over such contracts.<sup>16</sup>

#### **Electoral institutions**

The formal rules governing electoral processes, as well as the norms and other informal rules governing the way in which politics is conducted, also affect the political environment. These rules shape political incentives, and also the extent to which political parties have incentives for long-term cooperation.

New Zealand’s mixed-member proportional (MMP) style of elections affects the conduct of politics. In particular, since the party vote determines the allocation of seats in Parliament, this means political parties give priority to party “brand” issues over local electorate issues, and favour policies with appeal to the maximum of voters nationwide. This is to be contrasted with New Zealand’s former first-past-the-post (FPP) system, which required greater political attention to local electorate issues. This was particularly the case in marginal electorates that had the potential to determine which party would be able to govern.

The political focus on party brand issues under MMP has various consequences. It gives the Prime Minister strong incentives to protect the party brand, and results in presidential-like status being afforded the Prime Minister. All other things being equal, MMP electoral systems are empirically found to be associated with

<sup>16</sup> For example, public purchase in New Zealand generally requires the use of competitive processes such as tendering when the maximum estimated total value of the goods or services to be purchased equals or exceeds \$100 000 excluding GST (Ministry of Business, Innovation and Employment, 2014).

more fragmented parliaments, greater social spending, and larger deficits (Persson, Roland & Tabellini, 2007). Based on theory and evidence, they are also more likely to deliver public goods than under FPP (Menocal, 2011).

An important reason for why MMP electoral systems have these features is that they enable smaller parties to capture a greater share of votes. This means coalitions of parties are generally required to govern, necessitating political compromises, and catering more to the preferences of special interest groups than is the case under FPP. As a result, governments become more responsive to current issues of relevance to such interest groups. This is to shore up support from political partners, and to reduce the risks of third-party opportunism (such as attacks by opposition political parties).

The delivery of social services is affected by the process in which social policy is developed. This process involves gaining the cooperation of interested groups – including opposition political parties – particularly when programmes are long term. If the political process enables political cooperation, this facilitates efficient policies that can be adapted to changing circumstances over time (Spiller, Stein & Tommasi, 2003). Absent such cooperation, governments will implement policies using safeguards such as rigid policy rules, sacrificing efficiency and adaptability as a consequence. Alternatively, if such safeguards are too costly to implement, then political shocks are required to induce policy changes.

Spiller, Stein and Tommasi (2003) identify factors conducive to political cooperation. These include having a smaller number of political players, and players that interact repeatedly over time. Both of these make it easier to find and sustain mutually agreeable policy combinations. Delegating policy development to an independent bureaucracy also facilitates political cooperation, particularly when the preferences of the bureaucracy are located between those of opposing politicians. Conversely, political exchanges that are difficult to monitor can make it more difficult to sustain long-term political cooperation. Likewise, long-term cooperation is hard to sustain if the short-term rewards to opposing the Government are large.

### Soft budget constraint

Government providers of social services must compete for scarce public funds, but they do not share the risk of liquidation that face under-performing, non-government, service providers. This means that government providers are more likely to be allocated additional funds than be shut down in the event they are not financially viable, compared with non-government providers that face strict budget constraints. The resulting “soft budget constraint” softens performance incentives for government providers. This has negative implications for the delivery of cost-effective social services. It also has the potential to relieve the problem that giving strong incentives for a measurable activity (such as cost-saving) can result in the sacrifice of harder-to-measure activities (such as quality enhancement). This so-called *quality shading* is an important theme in the delivery of social services (Box F.2).

#### Box F.2 Quality shading

*Quality shading* can arise when a party pursues multiple activities, with quality affected by one or more activities in a way that cannot be contracted for. One reason is because quality-enhancing activities are hard to observe or verify. In that case, giving that party strong incentives to pursue one, contractible activity, can cause that party to reduce activity on the non-contractible activity (ie, quality). For example, giving employee incentive to a social services provider to reduce contractible costs might cause that employee to reduce their efforts in improving quality, if such quality is both costly to the employee and non-contractible.

Quality shading is an important concern for the delivery of social services if quality is important, quality-related effort is hard to measure or control, and if strong incentives are given to other activities with negative consequences for quality-enhancing activities.

Quality can mean different things to different parties involved in the delivery of social services. Governments are often concerned about the integrity of processes used to deliver social services. Social services providers might be concerned about employee competence. And consumers might value friendliness and approachability. In each case, quality shading can result – to differing degrees

and with differing consequences – if these dimensions of quality are hard to measure and other, easier-to-measure service dimensions are given strong incentives.

In some circumstances, quality-related effort might be *enhanced* (rather than reduced) when strong incentives are provided for other, more easily observed activities. Whether one type of effort increases or reduces the costs of providing another type of effort is a key consideration. Further, quality shading problems can be relieved in various ways, such as when parties are concerned about their reputation (which affects their future rewards from providing quality). Both cases are explored in section F.3.

*Source:* Holmstrom and Milgrom, 1991; Productivity Commission.

## Public sector employment conditions

Public sector employment conditions are another important set of rules affecting public sector performance incentives, and therefore the relative desirability of different forms of delivering social services. Standardised and rigid pay-scales, with constraints on both pay levels and performance-related pay, serve to limit monetary incentives in the public sector, creating both benefits and costs. However, non-monetary remuneration such as job security, equal employment opportunities, and an opportunity to easily move within the government labour market, compensate to some degree for lower monetary rewards than in the private sector. An important example of such non-monetary remuneration is the extent to which the public service ethos attracts motivated employees – that is, those with a shared sense of mission orientation (see Box F.5). This provides non-monetary incentives for performance, and substitutes to some degree for any lack of competition in the delivery of public services.<sup>17</sup>

## Social norms

Importantly, social norms are another important institutional constraint on the delivery of social services. This is particularly so in New Zealand given the greater priority over recent decades given to the partnership relationship established between Māori and the Crown under the Treaty of Waitangi. As a result, Māori values and norms – as well as Māori social contexts – have been increasingly recognised by social services providers. This necessarily involves greater diversity in the types of and provision models for social services, which allows them to be better-tailored to their target clients. For similar reasons, social services programmes now reflect other cultural norms and values to a much greater degree (eg, of New Zealand's Pasifika, Asian and other communities).

## Evaluating institutional changes

How institutional constraints arise or change is important when comparing the merits or shortcomings of different arrangements.<sup>18</sup> To some extent institutions are inherited, and therefore reflect historical accident as much as design. For example, that New Zealand has a legal system based on the common law arose due to it being a former colony of Great Britain. However, countries can adapt their “institutional endowments” as circumstances (eg, social services delivery technologies) and preferences (eg, biculturalism) change. Indeed, ultimately all institutions are amenable to adaptation. However, as “rules of the game” they tend to be adapted more slowly than the “games” played out within the confines of those rules.

Richter (2005) observes that “institutional change comes about through ‘entrepreneurs in political and economic organizations’ who realize that they could do better by changing the institutional framework” (p. 18). Indeed, retaining an ability to modify institutions is desirable for “adaptive efficiency” (North, 2003), allowing for the maximum of institutional choices and room to experiment, while providing institutions enabling failed solutions to be rejected. A criterion for judging the efficiency of an existing institution is provided in Box F.3. This is sometimes referred to as “comparative institutional analysis”.

<sup>17</sup> This observation also applies to NFP providers, with differences of degree. For example, NFPs often rely on voluntary labour and offer lower monetary incentives than government. In compensation, they rely more heavily than government on mission orientation to provide incentives.

<sup>18</sup> This discussion is based on that in Evans and Meade (2005).

**Box F.3 “Remediableness” criterion for judging whether an institution is efficient**

To judge whether or not an existing institution is efficient, Williamson (2000) notes that they should not be compared against hypothetical ideals. He instead proposes the following “remediableness criterion”:

...an extant mode or organization for which no superior feasible alternative can be described and implemented with expected net gains is presumed to be efficient. ...analysts can no longer condemn extant modes because they deviate from a hypothetical ideal... (p. 601)

This criterion “presses the public policy analyst to display a superior feasible alternative”, allowing also for any costs of implementation in the net benefit calculation. According to this criterion, an existing institution is inefficient only if no better alternative is feasible and can be implemented at greater benefit than cost.

*Source:* Evans and Meade, 2005; Productivity Commission.

### F.3 The make or buy decision

Whether goods or services should be bought in markets, produced within hierarchies such as governments, or through some hybrid organisational form, their exchange will be subject to some form of contract.<sup>19</sup> These contracts can be either short term (ie, spot market), or longer term (such as employment contracts and tendered contracts for the supply of services). They can also be *relational contracts*, with self-enforcing and relatively informal agreements taking the place of formal contracts and third-party enforcement (Baker, Gibbons & Murphy, 2002).<sup>20</sup>

The economics literature addressing the make or buy decision, and the more specific topic of government purchase, offers insights as to how exchange should be organised in the context of providing social services. This literature considers relative transaction costs, as well as the impacts on incentives of asset ownership and of contractual form. Asset ownership affects incentives when contracts relating to an exchange of goods or services cannot be fully specified in advance, with contractual incompleteness leading to hold-up risk (Box F.1). Incentive problems also arise when parties to the exchange have various forms of private information (Box F.6), even when contracts relating to the exchange can be fully specified.

An emerging economics literature also provides insights on the impacts of different organisational forms (ie, NFPs and FPs) in delivering social services. The literature involving economies of scale (in particular, in relation to hierarchies) and task bundling (including economies of scope) is also relevant.

This literature provides useful yet incomplete insights on the make or buy decision as it relates to social services. As stressed in section F.1, social services vary widely, depending on the client circumstances, the nature of providers, and how services are bundled. Social services delivery therefore needs to be considered on a case-by-case basis. As a consequence, each of the insights below will apply to varying degrees in relation to any given social service or bundle of services.

## Transaction costs and investment hold-up with incomplete contracts

### Boundary of the firm

Analysis of the make or buy decision began with a consideration of the boundary of the firm – that is, why there are firms at all, and not just markets (Coase, 1937). Under this analysis, firms exist – that is, they make rather than buy a good or service – when the transaction costs of purchasing goods or services on external (short-term – ie, spot) markets exceed those of organising internal production. However, while helpful, this explanation remains too general to be of much use in particular circumstances. It also fails to explain why firms do not just “outsource” production by contracting with some other firm instead of either buying the good or service on spot markets or making them internally (Williamson, 2002). Since this question is of

<sup>19</sup> Such contracts are often standardised and/or implicit, and do not need to be negotiated for every transaction.

<sup>20</sup> Relational contracts can be sustained when the long-run gains to maintaining a relationship outweigh the short-term gains from breaking it.

particular relevance to the delivery of social services – that is, whether government should provide those services or purchase them from private firms (rather than on spot markets), other explanations are required.

Coase's early explanation was expanded by considering how best to organise the production of a good or service when multiple workers are involved and the impact of their individual efforts to the final good or service cannot be well-measured (Alchian & Demsetz, 1972). Under this explanation, firms exist because specialist supervisors within firms substitute for the lack of market signals that can be used to assess the performance of individual workers. This means that firms provide useful monitoring technologies and therefore incentives in the context of joint production by multiple workers, where any one worker can free-ride on the efforts of others.<sup>21</sup>

### Hold-up risk

Further advances were made with the recognition that the hold-up risks facing investors in specific, particularly long-lived assets, affect how an exchange should be governed (Williamson, 2002; Klein, Crawford & Alchian, 1978; Box F.1). These risks are driven by the impossibility of fully specifying contracts in advance of investment, covering all future contingencies affecting the investment's value (particularly over longer horizons). They also arise when contracts cannot be perfectly and costlessly enforced, for example, because contract laws or the judiciary are weak, or the courts are too slow or expensive for effective contract enforcement. In each case the relevant contracts are said to be *incomplete*.

Before making relationship- or use-specific investments, the investing party has a greater degree of bargaining power relative to the other party to the relevant economic exchange. However, once the investment has been made, both parties can suffer from "bilateral dependency" – that is, they become "locked in" to the relationship to some degree. The non-investing party faces costs in trying to secure investment from an alternative provider once the contract has been entered into. Also, it faces the risk that the investing party acts opportunistically post-investment, for example by demanding higher prices before it agrees to further investments.

Conversely, the investing party faces a loss in asset value if the relationship is terminated. Because of this, the non-investing party may choose to act opportunistically once the investing party has made its investment. For example, they may choose not to fully pay the price agreed in the contract. Alternatively, they might simply choose to re-negotiate the contract if circumstances change post-investment (eg, if demand for their goods or services prove to be less than that expected when they initially contracted for the investment). Either way, the investing party can be "held up" by the non-investing party, which enjoys increased bargaining power post-investment.

Therefore, the value of the investment to the investing party can be undermined, post-investment, due to contractual incompleteness. Rational investors anticipating such hold-up risk will under-invest relative to the level that yields the best outcome for all parties. As a consequence, the parties can find it preferable to *vertically integrate*. That means that the party at risk of being held up owns the counterparty that supplies it ("make") instead of contracting with that counterparty for supply ("buy"). In this way hold-up risk is removed, the bargaining power of the investing party is not undermined post-investment, and a more efficient level of investment can be achieved. In the social services context, such vertical integration might be constrained by the ownership form of some providers, particularly NFPs.

### Hold-up risk when contracting with government

Two forms of opportunism giving rise to hold-up risk arise when NGOs making investments contract with governments (Spiller, 2008). Specifically, non-government investors face a general risk of government opportunism when governments cannot commit not to re-negotiate contract terms in light of changing information or circumstances (such as a change of government). This commitment risk is amplified where governments face little constraint on their decisions, such as when there is no formal constitution, or the judiciary is not independent and will not confront the Government.

Additionally, non-government investors contracting with government also face the risk of third-party opportunism. Interest groups and/or opposition political parties will selectively seek to pressure government

<sup>21</sup> This is called a problem of "moral hazard in teams".

when circumstances change so as to make earlier contracting decisions appear unwise, inappropriate or improper. Spiller argues that the risk of such third-party opportunism induces both governments and their contracting counterparties to favour more prescriptive contracts, and contracts with lower-powered incentives (see Box F.4). For example, they will favour “cost plus” contracts based on verifiable and therefore more defensible costs as means of reducing re-negotiation risks.

#### Box F.4 High-powered incentives and low-powered incentives

In this appendix the terms *high-powered incentives* and *low-powered incentives* are used for consistency with the relevant economics literature. Here they refer to the power of *monetary* incentives provided by a principal to an agent.

Incentives are “high-powered” (“low-powered”) when an agent receives a large (small) share of some risky outcome that is affected by their efforts. A simple example is a profit-related bonus in an FP, which the firm’s employee might affect through their work-related efforts (eg, to achieve cost reductions or sales).

However, whether an incentive is high-powered or low-powered also depends on its context. For example, an organisation providing services under a contract with the Government might face stronger overall incentives if their viability was threatened by losing that contract. Conversely, another organisation with the same type of contract might find that it offers relatively low-powered incentives. This could be the case if that organisation has a portfolio of contracts or greater financial reserves. In such cases, that organisation faces less severe consequences if it loses the contract.

Monetary incentives are to be distinguished from incentives provided by non-monetary means. An example of such an incentive is mission orientation (see Box F.5). When an agent is provided both monetary and non-monetary incentives, the relative importance of each will depend on the particular circumstances. Therefore, monetary incentives might be more important in some settings, and vice versa in others. Indeed, both types of incentive might arise in the same setting, but are applied to different kinds of agents (ie, those with and without strong mission orientation).

Further, monetary and non-monetary incentives can be complementary. This means that one form of incentive reinforces the other. However, it is also possible that they are substitutes. In that case, providing strong monetary incentives might reduce the power of (ie, “crowd out”) non-monetary incentives. These issues are explored in section F.4.

While the resulting contracts used to address third-party opportunism might be regarded as inefficient compared with private sector contracts, Spiller stresses that they address different contracting issues to those confronting non-government parties. This means that they are only inefficient under the remediableness criterion of Williamson (2000) if no other solution to third-party opportunism can be feasibly implemented at a lower cost (Box F.3).

### Resolving hold-up risk through ownership

However, just as Coase’s transaction-cost based argument can be criticised for not explaining how to choose between internal production and contracting for external production, so too can these hold-up based explanations of firm boundaries (Grossman & Hart, 1986). Further, the hold-up based explanations for firm boundaries can be taken to imply that vertical integration is never inferior to non-integration. This is because it recognises a benefit to vertical integration, but fails to account for the costs of integration, such as how integration reduces performance incentives for the acquired party.

Consequently, a literature based on *property rights* (ie, ownership) developed to account for this added dimension (Grossman & Hart, 1986; Hart & Moore, 1990). Again contractual incompleteness creates hold-up risks for parties making specific investments, with ownership – that is, residual rights of control – “completing” the relevant contract. Now, however, a trade-off is recognised. On the one hand, vertical

integration – that is, ownership of one party by the other – restores investment incentives otherwise damaged by hold-up risk. On the other hand, once the investing party acquires ownership of its counterparty, this reduces the performance incentives of the counterparty when making post-investment decisions regarding production.

Recognising this trade-off between incentives for investment and those for production, investments should be owned by the party making the investment that is most important for the value generated by all parties to production. However, this is provided that the benefits of ownership in terms of reduced hold-up risk outweigh the costs of reduced production incentives for the party that becomes owned.

An important example analysed by Grossman and Hart is the ownership of client files. Specifically, they examine whether client files in different types of insurance markets should be owned by insurance companies or insurance agents. Life insurance contracts are long-lived and rely to a significant degree on pre-contract screening. This suggests that insurance firms should own the relevant client lists. By contrast, general insurance contracts are short-term and post-contract service is relatively more important, favouring list ownership by insurance agents. This suggests that where the delivery of social services requires long-lived investment in customer-specific relationships, this favours information on those customers being owned by providers.

In the context of social services, such ownership-based explanations for whether services should be made or bought by government are relevant to the extent that the Government and non-government providers are required to make specific investments that may not be effectively governed by contracts. In principle, this will always be true to some extent, such as when either party requires specialist staff, information systems and other infrastructure. Exactly how important such investments are – and any resulting risks of opportunism and hold-up – need to be assessed in each particular case.

### **Different ownership types affect incentives**

An extension of this property-rights based literature – and the first to explore issues more directly relevant to the make or buy decision as it relates to social services – takes ownership assignments as given. It instead explores how the assignment of management rights to different types of owners (ie, government and FPs) affects their incentives (Hart, Shleifer & Vishny, 1997; hereafter “HSV”). The question HSV address is whether the *management* of a prison that is *owned* by government should be kept within government or contracted out to a private, profit-seeking firm. The manager, whether public or private, can make non-contractible investments in both quality or cost reduction, with the latter reducing quality. This setup illustrates a theme recurring in incentives literature, namely the risk of quality shading when agents pursue multiple tasks including quality, and quality is relatively hard to measure (Box F.2). Here it arises because the manager can make investments with conflicting impacts on quality.

In HSV’s model, government cannot commit to allowing a public manager to enjoy the benefits of any investments he or she makes. In practice, for example, due to the operation of budgetary processes or otherwise, if the manager makes investments in cost savings, that might simply lead to a reduction in future budget allocations.<sup>22</sup> Likewise, any rewards for improved quality are likely to be small for a public employee. As a consequence, the public manager faces little incentive to invest in either quality or cost savings. By contrast, the private manager faces stronger incentives for each type of investment. Whether the contract should be awarded to a public or private manager then hinges on whether providing stronger but conflicting incentives to the private manager for cost and quality is, on balance, better or worse than offering only weak incentives for either (as for the public manager). This contrast could be particularly important in the context of delivering social services – where cost is an important consideration, but so too is quality.

### **Mission orientation as a non-monetary incentive**

While the HSV approach moves closer to explaining when social services are better provided by government, or when they are better provided by private parties, it does not consider services in which there might be strong externalities or public good attributes often associated with such services. Besley and Ghatak (2001; 2003; 2005) present related models based on property rights, but for the provision of a public

<sup>22</sup> Analogous evidence for this in the context of the US health sector is provided in Duggan (2000). When US hospitals received subsidies to encourage service provision to low-income patients, they received no actual funding increase due to budget processes simply re-allocating their other funds.

rather than private good or service. They compare three forms of provision, emphasising how incentives for producing quality are affected by aligning the “mission orientation” of both consumers and employees (see Box F.5).

#### Box F.5 **Mission orientation**

A common theme emphasised in the provision of goods or services by either government or NFP providers is that of mission orientation alignment. *Mission orientation* refers to the values and objectives of organisations, workers or social services clients – relating not just to the exchange of goods or services, but also their availability to clients, and the means by which they are produced. Examples include the “public service ethic”, religious or otherwise-motivated “charity”, “sustainability”, “fair trade” and “equal opportunity employment”.

When the mission orientation of an organisation aligns with that of their employees, those employees receive on-the-job satisfaction that can substitute for monetary remuneration to some degree. This offers both improved incentives and reduced labour costs. Indeed, offering low wages can be used by organisations to attract only employment candidates with the desired mission orientation.<sup>23</sup>

When the mission of an organisation aligns with that of their clients, those clients enjoy benefits from consuming its goods or services over and above those offered by the same attributes of goods and services provided by other organisations that have different mission orientations. For example, a Māori client requiring employment services might prefer an iwi-based provider due to its better understanding of, and alignment with, that client’s cultural values.

Mission alignment enables an organisation to charge a premium for its goods or services, or to enjoy a larger market share than its rivals at a given price. However, when clients are poorly informed about an organisation’s mission orientation, this limits their ability to self-select into organisations that share their own alignment (section F.6). Likewise, it can be difficult for an organisation to discover the mission orientation of clients.

Changes in mission orientation can be organisationally disruptive (eg, causing a loss in morale). This means that organisations with significant mission orientation can be relatively rigid and conservative, adapting only slowly to changing needs or circumstances when this requires changing the mission. Mission can also be costly on other dimensions. For example, a mission that emphasises urgency militates against the accumulation of assets. Similarly, one that requires the NFP organisational form will constrain access to capital.

In the context of goods or services produced in non-market environments – for example, social services with public good attributes – competition to align the mission orientation of employers, employees and customers can substitute for the price- and quality-based competition and resulting incentives that are characteristic of exchange in private goods markets. In this regard, the mission orientation of government must be compared with that of NFP and FP non-government service providers.

*Source:* Besley and Ghatak, 2003; 2005; Productivity Commission.

Indeed, Besley and Ghatak stress that the better alignment of mission orientation not only improves incentives for performance and provides direct consumer benefits. It also plays an important role in the self-selection of employees and customers into organisations that match their own mission orientation. In other words, an organisation’s mission orientation helps it to better serve clients that share that orientation. It also helps the organisation to recruit employees who share that orientation. This reduces the monetary costs of hiring, and attracts employees who are better suited to the organisation’s activities.

<sup>23</sup> See Bénabou and Tirole (2013) for a discussion of the related theme of intrinsic motivation.

In Besley and Ghatak's setup, traditional, centralised public service provision (see Figure F.1(a)) is characterised as having a centralised mission orientation, rather than tailored mission orientations for different types of employees and customers. There is no competition with other providers for either employees or customers, since government is a monopoly provider in this case. Likewise there are only weak incentives for quality public service provision, since mission orientation alignment in this model arises in general terms (eg, public service ethos) rather than specifically.

Besley and Ghatak contrast this traditional model with one involving pure market provision. In this case, mission orientation is assumed to play a much more limited role, with high-powered monetary incentives used for employees when good performance measures exist. Instead of exploiting the motivation of employees with shared mission orientation, output market competition and the ability of employers to benchmark the performance of employees against that of others are used in conjunction with monetary incentives to ensure good performance. While shared mission orientation plays little role in providing client satisfaction, competition to meet client preferences improves satisfaction relative to the traditional, centralised model.

Besley and Ghatak offer their preferred approach to public service delivery – what they call “decentralized provision”. The hallmark of this approach is that multiple organisations with differing mission orientations vie to each attract social service clients having the same orientations as their own. Simultaneously, these organisations offer lower monetary incentives and instead vie to attract employees with shared mission orientations. Competition is now primarily in aligning the mission orientations of all three types of party – employers, employees and clients – and predicted to provide both optimal incentives and client choice in public services.

Returning to the ownership dimension of their extension of property rights models to public services, Besley and Ghatak conclude that ownership should no longer go to the party making the investment that makes the greatest contribution to the value of the service being produced, as is the case for private services. Instead, it should go to the party that more highly values the investment's benefits. This is because with public services (or those with significant and widespread external benefits) all parties value the services being provided, even if they are not directly involved in producing them. In the social services context, this suggests the provision of social services is best made by the party that best internalises the social benefits of those services. Possibilities include central or local government, NFPs (given their mission orientation), and local communities or iwi organisations.

### **Private choices of not-for-profit and for-profit status**

Glaeser and Shleifer (2001) extend the logic of Besley and Ghatak's preferred model by focusing on the incentives of NGOs to be set up with NFP instead of FP status. Again the context is one in which quality shading is of concern, and it is assumed that the parties creating these organisations are concerned about quality for reputational or altruistic reasons. NFP status is associated with capital non-distribution constraints, meaning operating profits cannot be used to provide incentives to employees via profit-related bonuses. Instead, employees of NFPs are restricted to non-cash performance-related employment (which is assumed to be valued less by employees than cash). Also, profit non-distribution constraints mean weaker incentives to maximise profits, and therefore to reduce costs. So they are also associated with reduced incentives for quality shading.

In this setup, NGOs are created with NFP instead of FP status if the benefits of committing to softer incentives and so reduced quality shading outweigh the costs of employees having to take net rewards in non-cash forms.<sup>24</sup>

### **Committing to low-powered incentives**

Acemoglu, Kremer and Mian (2007) further explore how different types of organisation – this time including government – can reduce quality shading issues by committing to use only low-powered incentives. For

<sup>24</sup> Relatedly, Bennett, Iossa and Legrenzi (2003) also analyse commercial NFP organisations in public service provision. Specifically, they observe the increasing use in the United States and United Kingdom of non-profit firms as new mode of service provision. Such firms rely on fees rather than donations (the more traditional source of NFP funding), and often operate in natural monopolies. Altruism and the use of voluntary labour play more limited roles than in traditional NFPs, while managerial incentives feature more prominently. These authors find that the reduced focus on profits in such organisations means they have less incentive to reduce quality to save costs, but question the effectiveness of their governance.

example, private firms can shield their operations from the high-powered incentives associated with markets by placing employees in teams (inducing the free-riding problems highlighted by Alchian & Demsetz, 1972). Indeed, governments can further commit to lower-power incentives. In this setup, weak incentives – whether in NGOs or government – not only alleviate possible quality shading issues. They also signal to employees with mission orientation that they face less conflict between orientation and incentives, which would arise to a greater degree in organisations focusing just on incentives.

These models suggest that alignment of mission orientation, between organisations and their employees, substitutes to some degree for high-powered incentives. Organisations therefore trade off high-powered incentives for performance against a reduced risk of quality shading, and a greater reliance on employee motivation to ensure performance.<sup>25</sup>

## Incentive issues under complete contracts with private information

This subsection considers providing incentives when contracts can be completely specified.

In this context, incentive problems arise due to parties having private information (rather than hold-up risks). These incentive problems fall into two broad groups: moral hazard and adverse selection (Box F.6). Simple versions of these two groups of incentive problem are presented, and then extended for the case of moral hazard.<sup>26</sup>

### Box F.6 Incentive problems due to private information – adverse selection and moral hazard

Incentive problems arise when parties to a contract have private information, with the nature of the incentive problem reflecting the nature of the private information. The most common examples are moral hazard and adverse selection.

- *Moral hazard* arises when an agent can take hidden actions. For example, the principal might be a health insurer, with the agent a customer who can take actions affecting their health risk which the insurer cannot observe. Alternatively, the principal could be an employer, and the agent an employee. The employee can take a number of actions that each affects outcomes that the employer cares about. However, they do so in a way that the employer cannot infer the employee's actions from those outcomes (eg, because random events or the actions of others affect outcomes).
- *Adverse selection* arises when a principal is contracting with an agent who has private information regarding their "type". For example, the agent might be a worker and the principal an employer, with the agent's type being their productivity. Alternatively, the principal might be a government department running a tender for services, with agents being possible service providers who privately observe their cost of service provision.

Another example of adverse selection involves a health insurance company as principal, and customers as agents. Each customer privately knows their health status, and therefore whether they will be costly to insure. If the insurer offers a flat-rate contract, possibly only costly customers will sign up. The insurer will then find it unprofitable to serve that market. Conversely, if the insurer offers different priced-schemes (eg, using different levels of co-payment or *excess*), then low-cost customers will self-select into contracts with higher co-payments. This helps the insurer to ensure that it appropriately prices contracts with low co-payments to reflect their greater use by high-cost customers.

*Source:* Dixit, 2002; Bolton and Dewatripont, 2005; Laffont and Martimort, 2002.

## Simple moral hazard

The simplest example of moral hazard involves a single principal seeking to contract for a single period with a single agent who can take a hidden action, for which the agent bears a private effort cost. While the action

<sup>25</sup> Indeed, it is possible that employee motivation is sufficiently strong in specific cases that it provides at least as strong an incentive as a high-powered monetary incentive. In that case, an organisation might enjoy both high performance (eg, cost savings) as well as low quality-shading.

<sup>26</sup> Moral hazard problems – particularly those involving multi-tasking and either multiple agents or principals – are of particular relevance when analysing social services. Also, while extensions of the basic adverse selection model are also relevant, the literature in this area is less developed than for moral hazard (Dixit, 2002).

of the agent cannot be verified, the result of the agent's action is observable. That result is affected by the agent's action in a probabilistic way, so it is impossible to infer the action from the result. Specifically, a greater level of effort by the agent increases the probability of a good outcome for the principal. An example could be an employer choosing the fixed salary and at-risk bonus compensation for an employee to deliver a service to a customer. The employee's actions are not sure to be successful, for example, because they depend on some random event such as whether a customer seeks the organisation's services.

Paying the agent only a fixed payment, irrespective of outcomes, removes all of the agent's risk. Assuming, for the moment, that the agent has no mission orientation (see Box F.5), it also removes all of the agent's incentives to choose an effort level that will deliver the principal's desired outcome. Conversely, paying only an at-risk component gives the agent maximum risk, as well as the strongest incentives for performance. In general, different combinations of fixed and at-risk payments give the agent different levels of risk and incentives.

The principal's problem is to set the terms of the agent's contract so as to maximise the principal's outcome net of the agent's reward. The principal does not have a free hand to do so, since it must offer the agent a contract that at least matches the agent's reward from not taking the contract at all and doing something else instead.<sup>27</sup> Also, the principal chooses the contract terms in the knowledge that the agent will choose the action that maximises the agent's payoff from the contract, given the contract terms it is offered.<sup>28</sup> Since the agent's actions are hidden, the principal instead bases the agent's payoff on the observable but uncertain outcome of the agent's actions. The agent's payoff is based on the risky payoff realised on the contract, net of private effort costs, including any discounting that the agent applies to net payoffs to account for their dislike of risk – that is, risk aversion.

If the agent's actions could in fact be observed by the principal, the incentive problem vanishes and the principal simply chooses the agent's payment scheme based on the agent's action (rather than the uncertain outcome of that action) to efficiently share the risk of the agent's activities, accounting for the relative risk aversion of the two parties.<sup>29</sup> This means that if the agent was risk-averse but the principal indifferent about risk, then the agent would be offered only a fixed, risk-free payment. In the reverse case, the agent would be offered only a performance-based reward. With both parties risk-averse, some intermediate risk-sharing is optimal.

However, when the agent takes hidden actions an incentive problem arises, so the principal's optimal choice of incentive power – that is, balance of fixed and at-risk reward for the agent – combines both incentive and risk-sharing components.

In the simplest such models, this means the agent's share of the risky outcome (eg, proportion of profits paid to the employee as a performance bonus) ranges from nil to 100%. Further, it increases as the agent becomes (relative to the principal) less risk-averse, or as the agent's private cost for each unit of effort falls. Lastly, the agent's share of the risky outcome decreases as the link between the agent's effort and the principal's outcome becomes more uncertain.

Indeed, if the agent was (relative to the principal) extremely risk-averse, the agent's private cost for each unit of effort was very high, and/or the agent could only very weakly affect the principal's outcomes through their actions, then the principal's best option would be to offer only very weak incentives to the agent (eg, through a predominantly fixed wage). Conversely, if the agent was (relatively) not risk-averse, faced only small private effort costs and could strongly affect outcomes through actions, then the principal would offer the agent a relatively high share of the risky payoff from their actions.

In the context of social services, high-powered financial incentives are created for non-government providers when they agree to deliver services of a specified quantity and quality at a fixed price. The costs of providing that quantity and quality can be uncertain, meaning the provider faces the risk that the agreed contract price is insufficient to cover those costs. Indeed, those costs can be affected by factors beyond the provider's

<sup>27</sup> This is known as the agent's participation constraint or individual rationality constraint.

<sup>28</sup> This is known as the agent's incentive compatibility constraint.

<sup>29</sup> This result is known as the Borch rule (Bolton & Dewatripont, 2005).

control, such as the nature or cooperativeness of the clients that they ultimately serve (the identities of which might be unknown at the time of contracting). A fixed-price contract effectively assigns 100% of the risky financial outcome to the provider.

If the provider has significant capital resources, then it will be in a better position to bear the financial risk created by a fixed-price contract than another provider that does not have those resources. This means that the incentives provided by such a contract might be appropriate for one but not the other of these providers, given their different degrees of risk aversion.

Alternatively, if the provider has a high degree of control over its costs, then the high-powered incentives provided by a fixed-price contract might suitably trade off the purchaser's desire for incentives and the risk-premium it must pay the provider to bear risk. However, if the provider has little control over its costs, then a fixed-price contract might impose too much risk on the provider and offer blunted incentives for performance.

In contrast to a fixed-price contract, a *cost-plus contract* in which the provider is assured of recovering all its delivery costs provides low-powered financial incentives. Indeed, the provider is assigned a 0% share of the risky financial outcome, and so the purchaser retains all the financial risk associated with the provider's delivery of the services. This might be justified if the purchaser is indifferent about risk (ie, is risk neutral), while the provider is highly risk-averse – for example, due to having limited capital resources. However, it would be inappropriate if the purchaser was highly risk-averse and the provider was risk-neutral. In that case, the provider more naturally bears financial risk, and so can optimally be offered stronger incentives (such as via a fixed-price contract). That would imply the provider optimally bearing more than 0% of the risky financial outcome.

### Simple adverse selection

In this case, the simplest example involves just one principal seeking to influence a single agent that can affect the principal's outcome. In this case, however, the agent does not take a hidden action, but instead is one of two or more possible "types" that the agent knows but the principal does not.<sup>30</sup> For example, the principal might be an employer, with the agent's types being whether or not they have a high productivity. While a low-productivity worker might have an incentive to misrepresent their type and claim to be high-productivity so as to secure a better wage, generally it is assumed that the reverse is not true (ie, a high-productivity worker would not claim to be low-productivity so as to get a low wage).

The challenge for the principal is to induce the agent to truthfully reveal their type. To do so, the principal chooses not just a single payment scheme based on the outcome produced by agent actions. Rather, in this situation the principal offers a *menu of contracts* to the agent and relies on the agent choosing the contract that is tailored to their actual type. Each menu option – one for each possible type of the agent – specifies the output the agent is to supply and the reward the agent will receive for that output.<sup>31</sup>

As for the moral hazard case, the principal does not have a free hand to choose the menu of contracts giving it the best payoff. It must also ensure that each agent type does better by contracting with the principal than not contracting with them. Likewise, the principal must ensure that the agent's payoff from choosing the menu option tailored to the agent's actual type cannot be bettered by choosing another menu option.<sup>32</sup>

The resulting menu of contracts that maximises the principal's outcome while respecting the relevant agent type's constraints has two clear features. First, the output demanded of the agent type that has an incentive to mimic the other type is not distorted relative to the output the principal would have chosen had it known

<sup>30</sup> This discussion relates to selection issues when agents are privately informed about their types. Another class of issues arising in this case relates to signalling of agent types (eg, by workers, to employers, via investments in education). For a discussion of this other class of issues, see Bolton and Dewatripont (2005) or Laffont and Martimort (2002). Attention is restricted to selection models as in Dixit (2002).

<sup>31</sup> That the incentive problem collapses to one of choosing one menu option for each agent type is a result of the so-called revelation principle (Bolton & Dewatripont, 2005).

<sup>32</sup> These are the analogues of the participation and incentive compatibility constraints arising in the moral hazard case. However, here there is now a participation constraint and an incentive compatibility constraint for each agent type. In practice only one agent type has an incentive to mimic the other type, so the principal needs only to respect the latter such constraint with respect to the agent that has this incentive. This follows from the application of the Spence-Mirrlees single-crossing condition (Bolton & Dewatripont, 2005).

that agent's type.<sup>33</sup> Second, the output of the type that might be mimicked is distorted downwards relative to the output the principal would have chosen had it known the agent's type, with the amount of the distortion reflecting the probability that the agent is the mimicking type. The reason for this distortion is to reduce the incentive for the potentially mimicking type to misrepresent its true type. However, in general the principal must still pay an "information rent" to the agent type that has the incentive to mimic the other type, to induce it to choose the menu option tailored to its true type.

Dixit (2002) gives an example in which the Government is the principal seeking to purchase a good or service from a private agent that can have either a high or low production cost that the principal cannot observe. Quality issues are not considered. The Government's problem is to design the menu of quantity-price contracts so as to ensure that the agent chooses to supply the Government irrespective of its true cost, and also chooses the contract tailored to its true cost. The Government does so to minimise its overall purchase costs. Since a low-cost producer might mimic the high-cost producer to secure a higher payment, the above discussion implies that the high-cost producer's quantity must be distorted downwards relative to the full-information case. Also, the price paid to the low-cost type is increased relative to the full-information case, reflecting the information rent the principal must leave to the potentially-mimicking type.

These insights are relevant to the discussion below regarding optimal purchase mechanisms, which also addresses the situation of multiple bidders that is each privately-informed about its supply cost. That discussion also considers optimal purchase when quality matters as well as cost, in situations where quality is possibly also private information of the bidders, and therefore in which quality shading is a possible issue.

The above trade-off highlights that in adverse selection problems the principal balances losing allocative efficiency (ie, choosing outputs for each agent type that reflect what could be achieved absent hidden information) against reducing the mimicking agent type's information rent – that is, a rent-efficiency trade-off. In contrast, the trade-off in the moral hazard case involves balancing the provision of efficient incentives to the agent against paying a risk-premium to compensate it for bearing the higher risk associated with higher incentives – that is, an insurance-efficiency trade-off.

### **Extensions of the simple models**

The above models involve only single principals, agents and time periods. In the context of providing or purchasing social services, it is important to also consider relevant extensions such as multiple tasks, multiple principals and/or agents, and multiple time periods. This includes discussions about contracting for outcomes versus contracting for inputs.

#### **Multiple tasks – quality shading revisited**

Holmstrom and Milgrom (1991) were the first to provide a formal, contract-based analysis of multi-tasking by a single agent in a situation of moral hazard. Specifically, a single principal is choosing the incentive contract of a single agent who can take multiple hidden actions which each affect the principal's observable outcomes in a probabilistic way. As in the simple moral hazard problem set out above, the agent bears a private cost in relation to each action. However, key to Holmstrom and Milgrom's approach is that the agent's level of one action can affect the private cost it bears from another.

For example, the agent's efforts to reduce the principal's observable production costs might make it privately more costly for the agent to take unobservable actions, increasing quality. In this case, the principal must weigh the benefits of giving the agent a strong incentive to perform the observable task against the possible reduction in the unobservable task that results. If greater observable effort raises the agent's private cost of unobservable effort, then a principal that cares about both effort types optimally responds by downward distorting the level of incentive offered for the observable task. While sacrificing performance incentives for the observable effort (eg, cost reduction), the principal preserves incentives for the other effort type (eg, quality production). This provides a more solid foundation for the quality-shading problem discussed in Box F.2.

However, undesirable trade-offs such as between cost reduction and quality shading do not necessarily arise under this approach. Key to the outcome is whether the agent's choice of one action level raises or lowers

<sup>33</sup> This result is known as "no distortion at the top".

their private costs of other actions. Indeed, if the private costs of each action type are independent of other actions, then the principal faces no trade-off at all, and can optimally choose incentive power for each action as in the simple moral hazard case. Alternatively, if the agent's choice of one action level *reduces* the private costs of other actions, then in fact the principal faces synergies rather than trade-offs between those tasks. Increasing incentives for an observable task might then induce the manager to increase effort on an unobservable task. A possible example of this is when a social service provider exerts effort to better organise case files, this makes it easier for that agent to undertake other tasks such as sharing those files with other colleagues.

Another application of multi-tasking is the case where agents can engage in *influence activities* (Dixit, 2002). Such activities might be inherently irrelevant to the principal's outcomes, but still provide useful information regarding other, hard-to-measure activities. In such cases, the principal may still find it desirable to offer incentives for influence activities. More generally, where agents engage in activities that are inherently irrelevant to the principal but which affect activities that are inherently relevant, then this too affects whether the principal seeks to encourage them or not. For example, employers might subsidise worker participation in social sports clubs to help foster positive work relationships. Alternatively, it might contractually prohibit employees from moonlighting or engaging in extra-mural activities that conflict with work activities or risk injuring the employer's reputation.

A challenge of job or organisational design is to try to exploit such synergies, or conversely reduce such trade-offs. This is particularly important if social services providers (whether government or private) collaborate across multiple organisations, or combine multiple tasks within their organisation. Depending on how some tasks affect the private costs of parties pursuing other tasks is important for determining whether bundling tasks increases or reduces the appropriate level of incentives for each task.<sup>34</sup> For example, combining social services that tend to be delivered to the same clients might involve synergies in terms of providers' costs. This is because providers can provide multiple services in single client engagements. Conversely, bundling services that involve delivery to different clients offers no such gain, and possibly increases providers' private effort costs in serving all their clients.

### **Contracting for outcomes versus contracting for inputs**

This discussion is particularly relevant to "contracting for outcomes", as well as "contracting for inputs". Box F.6 highlights that moral hazard problems arise when one or more of an agent's activities (eg, cost reduction) cannot be observed and therefore cannot be contracted for directly. This forces the principal to instead contract on observable variables – such as outcomes or inputs (such as hours worked). An advantage of contracting for inputs is that it is relatively objective and simple to implement. Conversely, while contracting for outcomes should in principle be preferred to contracting for inputs, outcomes can be harder to measure precisely or objectively.

In either case, contracting on any one measured variable, and providing incentives based on that variable, will naturally cause agents to focus on delivery in terms of that variable. At the same time, that can affect the agent's private costs of delivering on other variables – either in a supportive or undermining way. For example, providing incentives to work a set number of office hours each day might increase an agent's personal cost of engaging with clients outside office hours. This could detrimentally affect the agent's performance by restricting opportunities for getting to know clients' needs in more conducive settings.

Contracting directly for performance – that is, outcomes – might either increase or reduce inputs (eg, by leaving agents to choose how best to deliver outcomes). And if only some outcomes dimensions are contracted on, there is still the possibility that non-contracted outcome dimensions will be affected by how incentives are provided for in contracted outcomes. In short, agents can be provided with both explicit and implicit incentives. When undertaking multiple tasks (whether inputs or outcomes), explicit incentives for one affect the private costs of delivering others. In that sense, those explicit incentives also provide implicit incentives – either reinforcing or conflicting – for those other tasks. Accounting for these implicit incentives is an important challenge in contracting, job design, and organisational design problems.

---

<sup>34</sup> Clearly other considerations are also important, such as the governance of any collaborative activities.

### Multiple agents — moral hazard in teams, and relative performance measurement

As identified in the multi-tasking case, interactions can be key to affecting optimal incentives. This applies equally to the case of a single principal contracting with multiple agents, each of whom can take hidden actions on a task (or range of tasks) – referred to as “moral hazard in teams”. In the simplest case where there are no interactions among team members, the principal can design optimal incentive contracts (as in the simple moral hazard case). Otherwise the incentives for each agent must be determined jointly.

If the team is rewarded as a whole, then agents are able to free-ride on the efforts of others. If the tasks of each agent are complementary (one agent’s effort increases the output of other agents’ efforts), then the principal can overcome this by offering strong incentives to some team members because this then increases the output of others, which increases the rewards to them from exerting higher effort also. However, if the tasks are substitutes (one agent’s effort decreases the output of other agents’ efforts), then more uniform rewards are better (Winter, 2004).

With multiple agents it can be optimal for a principal to implement relative performance measurement for each agent – what is sometimes called benchmarking, yardstick competition, or tournaments. In this case, each agent is rewarded based on their performance relative to other agents. The advantage of this approach is that it nets out the random element of each agent’s performance that is common to all agents, which reduces the overall risk faced by each agent while highlighting their individual contribution to performance. In reducing the risk that agents face, the principal can then offer higher incentives than would otherwise be offered for a given risk, since the principal faces an efficiency-insurance trade-off in moral hazard problems as discussed above.

However, if agents’ outputs do not share a random influence, then relative performance rewards can instead undermine cooperation and even induce agents to sabotage the outputs of other agents. That way they improve their performance ranking relative to others. Indeed, in that situation, allowing agents to collude (ie, agree to work less intensively) could even be beneficial to the principal. While collusion involves sacrifice of team output, relative performance incentives can still be used. This is beneficial in situations where individualised incentives might require prohibitive risk-transfer to employees, and therefore are infeasible (Bolton & Dewatripont, 2002).

### Multiple principals – common agency or servants of two (or more) masters

In the above discussion of the distinctive features of social services, both multi-tasking and multiple principals were emphasised. The latter, in which an agent serves more than one principal, is known as *common agency*, and can be thought of as a servant serving more than one master. In such situations, it is little surprise that having multiple principals serves to weaken the incentives faced by the agent (Dixit, 2002). The reason is that each principal wishes to offer the agent positive incentives to pursue the activities it cares about, and negative incentives to pursue the activities cared about by other principals.

This dilution of incentives is reduced when the agent’s actions for each principal are complementary (ie, pursuing one action lowers the agent’s private costs of pursuing the other actions). However, it is worsened when the agent’s actions are substitutes (ie, one action increases the private cost of other actions). In fact, this offers a possible rationale for an agent’s different activities to each be ring-fenced and sheltered from the view of principals who are not directly affected by that activity (Dixit, 2002). Alternatively, principals could cooperate to agree on a jointly optimal incentive scheme for the agent.<sup>35</sup>

### Contracting over multiple time periods

All of the above models have focused on single-period contracting problems. In multi-period settings repeated interaction, learning, signalling and reputation-building all play roles in affecting the optimal level of incentives that a principal should set an agent (Dixit, 2002). For example, principals can offer smoothed payment schemes that reduce the agent’s risk over time while still preserving incentives (similar to the benefits of relative performance measurement discussed above). Conversely, implementing incentive

<sup>35</sup> This could be relevant, for example, in the case that a single government ministry is seeking to achieve outcomes valued by more than one minister. Cabinet then serves as a coordination point to ensure that the ministry is not provided with multiple, conflicting incentives by those different ministers.

schemes on a period-by-period basis when the principal cares about the agent's long-term performance can induce the agent to game the scheme, such as by slackening off when early rewards are good.

Indeed, the correct design of per-period incentives can be very important when outcomes such as quality are observed over time. For example, if quality shading is a concern, then the principal might over-compensate the agent for producing quality for so long as quality is maintained, but punish poor quality with a low payment when detected. However, an agent's concern for reputation can substitute for incentives in such situations. Specifically, if an agent expects to receive later rewards for good early performance, then this can avert poor early performance and resulting later problems. Indeed, agents may exert high effort levels early in their contractual relationship so as to signal to the principal that they have high ability, foregoing early rewards in the expectation of being rewarded in the future. Such reputational concerns can be an important constraint on FPs when quality shading is of concern (section F.4).

### Other incentive considerations

In addition to discussing extensions of the simple moral hazard problem, other special issues are relevant to considering incentive power in the context of the delivery of social services. These include a return to the earlier discussion of mission orientation (see Box F.5), as well as discussions of how incentives are affected by each of competition and different forms of ownership (ie, NFP, FP, government). Other mechanisms to better align the interests of principals and agents, such as monitoring and bonding, are also discussed.

### Mission orientation revisited – incentives, selection and crowding out

The alignment of mission orientation between principals and agents can substitute for explicit monetary rewards. This not only induces agents to participate in the principal's contract at a lower wage, but offers an alternative means of inducing desired effort levels. In particular, it can alleviate the need for high-powered incentives on observable outputs, which reduces possible quality shading issues where quality is not observable.

The related literature emphasises another important impact of incentives – their impact on self-selection by employees into employers. For example, increasing incentives to nurses can have the detriment of causing less vocation- (ie, mission-) oriented nurses to take nursing jobs, to the detriment of patients (Heyes, 2005; Le Grand, 2003).

Alternatively, by offering high-powered incentives, employers can attract (ie, *cherry-pick* or *cream-skim*) the most able employees whether mission-oriented or not (Dixit, 2002). In some cases this can serve to crowd out any mission orientation they might have (Bénabou & Tirole, 2003; Frey & Oberholzer-Gee, 1997).<sup>36</sup>

Conversely, self-selection by less-able employees into employers offering lower-powered incentives but greater alignment of mission orientation provides an alternative means of providing incentives, albeit for a cohort of employees of lesser ability. Comparisons such as these are relevant, for example, when comparing high private-sector pay levels against lower public-sector pay levels but greater public service ethic. In turn, the same comparisons can be made between relatively high pay levels, but more generic mission orientation in the public sector, against the even lower pay, but more specific mission orientation of non-government NFP organisations. Indeed, with NFPs often relying on volunteer workers, exclusive mission orientation in NFPs plays an analogous role to that of exclusively high-powered incentives in private-sector FPs.

Competition in the labour market for employees can affect both the level and incentive power of compensation, with resulting employee selection affecting intrinsic motivation (ie, mission orientation) – such as in relation to quality (Bénabou & Tirole, 2013). If an employer has market power in any given labour market, it is relatively free to offer lower wages and lower-powered incentives. This attracts employees with greater intrinsic motivation, which provides incentives for quality provision and alleviates quality-shading concerns. However, if that employer must compete with others for the available pool of employees, then this results in increased wage levels and higher incentives. The increased reliance on high-powered incentives

<sup>36</sup> Kosfield and von Siemens (2011) also show that monetary incentives induce performance but simultaneously shape the motivational characteristics of the employees attracted by those incentives. In particular, competition for employees can lead to employees with lower mission orientation working in organisations offering high monetary rewards, and featuring low levels of employee cooperation. Conversely, more mission-oriented employees choose to work in organisations offering lower monetary rewards, which also feature greater levels of cooperation between employees. Their approach explains why it is possible, for example, to have NFPs and FPs co-existing in the same area of activity, with different levels of cooperation in each.

attracts workers with less intrinsic motivation, potentially aggravating issues such as quality shading. Under this model it is possible, for example, that competition between government and non-government employers for social services workers could serve to undermine incentives provided by mission orientation.

### Competition and incentives

It is often presumed that greater competition – here in an organisation’s output market rather than the labour market – should generally improve agents’ performance incentives. For example, to survive in a competitive environment requires firms to better meet their customers’ preferences for variety and quality, as well as price. Alternatively, greater competition can improve incentives by making principals more informed about their agents’ actions (Hart, 1983), or because it increases the risk of bankruptcy, the threat of which induces higher managerial effort (Schmidt, 1997).

However, theory is ambiguous on this point, with increased competition possibly improving or worsening an agent’s actions (Beiner, Schmid & Wazenried, 2011; Holden & Posner, 2014). On the one hand, greater competition increases the potential reward to reducing costs (a business stealing effect). On the other hand, increased competition also reduces the scale of each operator, which reduces the rewards from improved performance (a scale effect). While the first effect encourages principals to provide agents with stronger incentives, the second effect does the reverse. Which of these two effects dominates depends on the form of competition (ie, number of firms, degree of product differentiation, etc.). Increased competition can therefore either increase or decrease principals’ optimal choice of incentive power. Given this ambiguity, it is not clear that competition between government and non-government providers of social services, or among non-government providers of social services, necessarily improves incentives. The impact of competition on incentives must be considered in the light of particular circumstances.

In respect of quality, competition can eliminate the price premium required to support its provision (Kranton, 2003). This can explain why industry associations – especially in professions – sometimes seek mechanisms to limit competition. These include entry barriers such as training and accreditation requirements. Once again, the impact of competition on incentives (here for quality) will depend on particular circumstances.

### Ownership form and incentives

Much of the literature on incentive problems under private information considers incentives within FPs, or between governments and such firms. There is also an emerging literature on the incentive issues of NFPs (eg, Bennetson, Iossa & Legrenzi, 2003; Bacchiega & Borzaga, 2001). However, and related to the discussion of competition and incentives, there is a clear gap in the contracting literature regarding how competition between different types of organisation – ie, government, FPs and NFPs – affects incentives.

Approaching the question of how ownership form affects incentives at a more general level, a number of considerations are relevant. First is the ability of different organisations to bear different kinds of risks, since the moral hazard discussion above highlights the efficiency-insurance trade-off in that case. Since FPs can accumulate capital with which to withstand financial losses, and shareholders can diversify their investment risks, FPs can be regarded as relatively less risk-averse, and so are able to bear relatively high incentives (which carry higher associated risks).

This is the case, for example, compared with NFPs, which cannot issue ownership claims such as shares to external investors and so can face more binding capital constraints. This reduces their ability to bear the risk of financial losses. Further, their often tight focus on a particular mission means they have little ability to diversify their financial risks within their own organisations. And their lack of tradable ownership claims complicates the formation of joint ventures or mergers with other organisations so as to share their risks. These factors mean that NFPs will have relatively high risk-aversion (as compared with FPs, all other things being equal) and so are less able to bear high-powered incentives, which necessarily involve risk transfer. NFPs must therefore rely more on mission orientation alignment than monetary incentives to induce good performance. Whether mission orientation alignment provides stronger or weaker performance incentives than monetary incentives will depend on the particular circumstances.

Conversely, government might be regarded as relatively less risk-averse given its large access to capital and limited risk of liquidation (due to the soft budget constraint discussed earlier). However, as emphasised by Spiller (1998), governments have high degrees of political risk-aversion. In the context of complete

contracting under moral hazard this suggests that a government should favour offering high-powered incentives if it purchases goods or services from non-government parties. However, Spiller's incomplete contracts framework suggests the opposite conclusion; that is, low-powered incentives as a means to reduce third-party opportunism.

The particular characteristics of government, FPs and NFPs can be regarded as constraints on the types of incentives that can be used. Conversely, they can be regarded as organisational adaptations designed to address incentive issues not remediable by other means. For example, both NFPs and governments can be seen as organisational forms committing to low-powered incentives, greater mission orientation, and less quality shading, than FPs. Likewise, the use of low-powered incentives, mission orientation, and relative performance measurement can be seen as responses to resolving moral hazard issues in teams. Ultimately, ownership form is a choice variable of the parties organising an activity (Glaeser & Shleifer, 2001). It can therefore be considered a complement to choices over other explicit or implicit incentives. As Dixit (2002) puts it:

...whether an activity is carried out in the public sector or the private sector is itself endogenous, responding to the differences of technologies of carrying out the necessary political transactions, namely coping with information asymmetries, making credible commitments, and enforcing contracts. (p. 717)

### Other performance mechanisms

In practice a range of mechanisms can be employed to encourage agents to perform in the interests of their principals. For example, Jensen (1993) identifies four types of "control forces" to relieve the "agency costs" that arise when agents' interests diverge from those of their principals.

- First, there are legal, political and regulatory control forces, such as company law and competition law. These provide general controls on the behaviour of an agent (eg, an FP manager).
- The second set of control forces include internal control systems, such as boards of directors, financial reporting requirements, and the use of external auditors. These emphasise the usefulness of introducing technologies to monitor agents' performance when principals lack the time or expertise to do so directly. Ensuring that the right monitoring is performed, and that monitors themselves have good incentives to monitor (and not to collude with) agents, are important challenges. Further, just as explicit monetary incentives can crowd out incentives provided by mission orientation, so too can monitoring conflict with agents' trust and loyalty (Frey, 1993). Greater monitoring can be demotivating for agents if they perceive it as a sign of distrust, particularly in close principal-agent relationships. However, this effect is less important, and therefore monitoring is more effective, in more impersonal settings such as markets.
- The third set of control forces includes competition in both output and input markets. The role of output market competition in providing incentives was discussed above. With respect to labour market competition, the role of mission orientation alignment in helping organisations to compete for staff has also been discussed, as have the incentives for employees to build good reputations through good performance in multi-period contexts. In either case *selection* plays an important role. This selection can be by customers or employees as to which organisations they use or work for. Alternatively, organisations can select which employees to hire, or which customers to serve. Either way, selection processes help to reduce misalignments between the interests of principals and agents. An example of how selection is achieved is through offering monetary incentives. As stressed above, the level and composition of such incentives can reinforce or undermine non-monetary incentives, such as mission orientation, by affecting the types of employee to which they appeal.
- Jensen's fourth set of control forces includes the *market for corporate control*. This refers to the process by which under-performing organisations are acquired by other organisations that then implement changes in strategy and output delivery, or inject capital, to improve performance. While this process can be more readily implemented in organisations with tradable ownership rights (eg, FPs), this is less the case for organisations lacking such tradable ownership (eg, NFPs and government departments).

Another important mechanism for aligning incentives is bonding (Jensen & Meckling, 1976). This can involve organisations using debt financing as a way to promise to pay cashflows to financiers (ie, principals) rather

than dissipating them on managerial benefits. Or it can involve agents' performance-related remuneration in one period being "banked" and only made available to them if good performance is sustained in later periods.

Therefore, a combination of mechanisms ultimately ensures that agents act in the interest of their principals. The mechanisms include explicit and implicit monetary incentives, non-monetary incentives, competition, monitoring, bonding, selection, and legal rules. Since it is costly to implement these mechanisms, and they are each imperfect to some degree, in practice they will not eliminate all divergences of interest between agents and principals.<sup>37</sup> In each situation the challenge is to find the best feasible set of arrangements, taking into account all relevant trade-offs.

## Hierarchies and economies of scale

As highlighted in the discussion of Figure F.1, the delivery or purchase of social services by government is typically very hierarchical. High-level decisions as to what social services to deliver and how they are to be purchased are made at Cabinet level, and then are mediated via multiple management levels until they reach the public or employees who actually provide those services at the customer level. In contrast, the non-government NFP or FP providers of social services enter into Figure F.1 at lower levels, and typically involve much less hierarchical organisations.

A full discussion of hierarchies – considering both coordination and incentives – is beyond the scope of this appendix.<sup>38</sup> However, certain considerations are of particular relevance. They are the importance and sources of scale economies in the delivery of social services, the control and therefore incentive characteristics of hierarchies, and the importance of information in hierarchies.

### Economies of scale

Hierarchies have historically been the hallmark of large organisations. Such organisations are common when there are large economies of scale in production, meaning that production costs for each unit are minimised only with large-scale production. They arise in particular when production requires large upfront investments (eg, in fixed assets). This is often associated with monopoly provision of a good or service, when sufficient scale can be achieved by only one firm.

One potential source of such scale economies in social services relates to IT. When the coordination of services requires large amounts of information, this can necessitate large investments in specialised IT infrastructure; in particular in high-cost software systems customised to particular problems or organisations. Efficient use of such investments requires large-scale production and organisations that are capable of sustaining that level of production.

### Span of control, loss of control and delegation

Hierarchies emerge as an organisational form capable of sustaining high levels of production, trading off the gains from scale economies against the costs of increasing span of control and loss of control as greater reliance is placed on delegation (Bolton & Dewatripont, 2005). As the workforce required to achieve high output levels expands, this increases the number of employees reporting to the organisation's principal – that is, increasing that principal's span of control. This reduces the ability of the principal to monitor and control the activities of any given employee, and gives rise to the need for the principal to delegate monitoring to mid-level managers. In so doing, the principal trades off the advantages of delegation against the loss in control and other disadvantages that this delegation entails.

The disadvantages of delegation include the principal being distanced from the grass-roots activities of the organisation, and therefore being less responsive to changes in its operating environment. However, at the same time, the principal gains the benefits of being able to coordinate activities over a great range of tasks, despite any losses from delegation. In terms of making project choices, delegation encourages risk-taking by juniors, while centralised decision making reduces the risk of errors (Sah & Stiglitz, 1986).

<sup>37</sup> In Jensen and Meckling's terminology, there will always be "residual agency costs".

<sup>38</sup> Relevant contributions are discussed in Hart and Moore (2005).

Another disadvantage is that delegation increases the risk of collusion among the principal's subordinates. For example, they might only selectively report to the principal, or agree to reduce effort levels while preserving relative performance. As discussed earlier, in some situations such collusion might in fact provide the principal with a means of preserving incentives among multiple agents when the actions of those agents are substitutes. However, in general collusion carries costs to the principal in terms of both worsened outcomes and the direct and indirect costs of introducing monitoring and verification technologies such as external audits. To reduce the risk of collusion, it is generally necessary to provide reduced incentives at lower levels of the organisation, as these reduce the possible gains to colluders (Dixit, 2002).

The advantages of delegation include the use of specialised managers to effectively monitor a greater number of subordinates on the principal's behalf. Indeed, such specialised managers may be better able than the principal to elicit performance from subordinates – with less information-rent transfer than in the simple moral hazard example discussed earlier (Dixit, 2002). Further, by delegating authority to lower-level subordinates a principal can economise on the need for explicit incentives, with agents deriving greater satisfaction from decisions over which they have greater control (Aghion & Tirole, 1997). The converse is also possible – that agents can feel demotivated when they have little control over decisions that affect them.

Delegating decisions to lower tiers in a hierarchy can also improve asset use by managers, while reducing the benefits of investment coordination across the organisation. This makes centralisation preferred when coordination gains are large, while decentralisation (and even organisational separation) is preferred otherwise (Hart & Moore, 2005). Delegating to an intermediary with preferences between those of the principal and the agent enables the intermediary to gain the agent's trust and foster initiatives to a greater degree than if the relationship was more direct (Dessein, 2002). This might be particularly relevant in the delivery of social services, in situations where social services clients might be unwilling to share information directly with government agencies. They might, however, be more willing to do so with non-government service providers whom they regard as more neutral or sympathetic.

### Information in hierarchies

An advantage of centralised decision making in hierarchies is the ability to coordinate information obtained across multiple activities. This comes at the cost, however, of decision makers being more removed from the organisation's environment. Conversely, an advantage of decentralised decision making is the ability to better identify and adapt to changing market conditions. This is particularly important when organisations need to innovate to survive in competitive environments (Guadalupe & Wulf, 2010).

Hierarchies can also enjoy certain informational advantages. One advantage is that they generate competing sources of information. Indeed, even when biased subordinates are used to make the case – that is, advocate – for their preferred projects, this can improve the informativeness of signals reaching the organisation's principal relative to having only one advocate (Dewatripont & Tirole, 1999). This is because it is possible to offer competing advocates incentives that more closely align with the project they advocate, while a single advocate of conflicting projects faces diminished incentives.<sup>39</sup> This means competing advocates have stronger incentives to gather information than a single advocate. Also, advocacy provides strong incentives to appeal unfavourable decisions, which further improves information provision to the principal.

Centralised decision making within hierarchies can be sustained by greater use of information standardisation such as specialised language or jargon. This includes the use of standardised expressions with widely understood meanings. Such standardisation balances the benefits of jargon that simplifies internal communication against those of more commonly used language that facilitates external communication (Cremer, Garicano & Prat, 2004).<sup>40</sup>

### Relevance to social services

This discussion of hierarchies suggests reasons for why government might remain a dominant provider of certain social services, even when competing non-government provision is possible. Government potentially

<sup>39</sup> This is analogous to the weakening of incentives that arises when an agent answers to multiple principals, as discussed earlier in this section.

<sup>40</sup> Indeed, the development of a single engineering database to coordinate subcontractors' activities proved instrumental to the development of the B-2 Stealth bomber.

has a comparative advantage relative to other providers in the large-scale provision of social services, involving significant economies of scale or centralised, high-level coordination. Further, government's use of standardised communications reduces some of the costs of hierarchical communications, and potentially offers benefits in terms of competing communications from parties (eg, ministries) with divergent biases.

Conversely, non-government providers can provide both bundled and more tailored services relative to centralised state provision. This requires those providers to be more decentralised and closer to market information – particularly in relation to their clients. Being so enables them to be more responsive than government to changing client requirements or preferences. Decentralisation also facilitates lower-level coordination across service areas. Further, non-government providers lack the resources of the state to be engaged in large-scale provision, particularly when they are NFP, and so cannot raise capital to the same degree as FP organisations. This removes a rationale for them organising more hierarchically.

Where government adopts standardised codes for communication, this provides a means for better integrating non-government providers into the government's coordination system. This better enables government to access the benefits of decentralised provision by non-government providers while retaining the potential benefits of centralised coordination.

## **Task bundling**

Section F.2 stressed that social services often need to be delivered in a bundle, reflecting the particular needs and circumstances of social services clients. This raises issues of coordination and economies of scope, as well as efficient incentive provision and ownership form.

### **Coordination**

The above discussion of hierarchies emphasises how centralised decision making can be useful in achieving coordination. However, this comes at the cost of distancing high-level decision makers from client-related information. This means there is a trade-off between top-down coordination across multiple activities and access to client information when choosing between hierarchical and decentralised organisation. This trade-off becomes less severe when technologies exist for coordinating among decentralised activities. Technologies for decentralised coordination enable task bundling decisions to occur closer to the client level.

### **Economies of scope**

Setting aside coordination and incentive issues, economies of scope represent a simple way to think about whether social services should be bundled or not. Economies of scope exist when combining two or more activities into a single organisation is less costly than producing them separately by specialised organisations (Panzar & Willig, 1981). Economies of scope arise when there are learning spill-overs (ie, learning in relation to one task helps to better deliver on another). They also arise when activities share a common input in a non-rivalrous way (ie, use of the input on one activity does not reduce its availability on another, such as client files). While economies of scope suggest a simple test for considering whether services should be bundled, the incentive impacts of bundling must also be taken into account.

### **Ownership form, delegation and incentives**

Combining tasks to provide complementary rather than conflicting incentives was discussed earlier in the context of multi-tasking under moral hazard. A key insight was that indirect incentives for unobservable activities can be provided by providing direct incentives for observable ones. However, this required that the agent's private costs of providing the unobserved activity were reduced when they undertook the observed activity. If that were not the case, then direct incentives on one task could provide conflicting indirect incentives on another.

Ownership form and task bundling can play an important role in providing incentives for cost, quality and "market orientation" (Holmstrom & Tirole, 1991). They consider an organisation with two divisions, each producing a component that can be combined to make a final product, or sold on outside markets. Bundling the two activities within the same organisation enables the managers of each division to be offered coordinated incentives, which can be used to encourage cooperation. An incentive problem arises because each manager can make unobservable investments in quality and cost reduction. They can also choose how

much market orientation to build into their component, reflecting the extent to which it is not specialised and so is suitable for external sale, rather than so specialised it can only be used within the organisation.

Holmstrom and Tirole's model predicts that decentralised bargaining between the two managers as to internal transfer prices is strongly affected by whether or not they are permitted to sell their component externally. If they can, then they choose an excessive level of market orientation to make their component attractive to outside buyers. This increases their bargaining position when agreeing transfer prices with the other manager. In turn this means they are better able to capture the benefits of quality investments, and so encourages quality provision.

However, when the managers are prohibited from selling their components externally, they cannot improve their internal bargaining position through their choice of market orientation. This comes at the cost of not being able to improve their returns from quality investments, so involves lower-quality provision. But it also comes at the benefit of improved internal coordination and cooperation. The optimal degree of decentralisation trades off the quality-related benefits of being able to trade externally and higher market orientation against those of improved coordination and cooperation when external trade is banned.

Another trade-off in this setup is between the coordination benefits of bundling the two tasks within the one organisation, versus the temptation for the organisation to try to influence the transfer pricing bargaining between the two managers.<sup>41</sup> Where that temptation is too great, unbundling the two tasks may be preferred.

In the context of bundling the delivery of social services, Holmstrom and Tirole's model suggests that bundled provision with no chance of unbundling may be preferable when quality provision is less important than coordination and cooperation across activities, and encouraging specialised investments (eg, in client relationships). Conversely, bundling with the opportunity to unbundle may be preferable when quality provision is relatively more important.

Unbundling may be preferable when it is not possible to fully delegate to providers of different services decisions as to the price they charge each other for those services. This may be more an issue in relation to government provision of social services than provision by either NFPs or FPs. It perhaps suggests another reason why government provision is often via "silos", and achieving client-level (rather than high-level) coordination is difficult. However, it also suggests that both coordination and specialised investments might be more supportable in NFPs and FPs.

## **Collective action as an alternative to government organisation**

Much of the economic analysis of organising the delivery of goods or services assumes individuals act rationally and in a self-interested manner. While this assumption works well in predicting behaviour in market settings, it works less well in predicting outcomes in situations involving collective action (eg, communities organising their own social services). In particular, there is considerable evidence to refute the proposition that self-interested individuals will not contribute to the provision of a public good unless the group from which they come is small, or they can be coerced to act in the collective interest of the group (Ostrom, 2000).

To better predict outcomes in collective action situations, it is necessary to assume other types of individuals to the standard "rational egoist". In particular, economists' understanding of collective action problems has been assisted by assuming the existence of "conditional cooperators" as well as "willing punishers" (Ostrom, 2000). The former are prepared to initiate cooperative actions when they trust that a sufficient number of others will reciprocate. Willing punishers can also be conditional cooperators, but they are willing to punish (or reward) individuals who attempt to free-ride on the contributions of others.

The existence of individuals who are not rational egoists is confirmed by experimental and field evidence (Ostrom, 2000). This evidence suggests that there are individuals who bring with them norms and values that sustain cooperation. Important among these are trust and reciprocity. However, in larger populations it can

---

<sup>41</sup> This is comparable to the problem inherent in government ownership identified by Hart, Shleifer and Vishny (1997). Specifically, government is less able than private firms to commit to leaving the returns from investments made by an employee to that employee.

be harder for individuals to learn whether other individuals share such norms. In such populations the trust and reciprocity required to sustain conditional cooperation can break down, leaving only rational egoists.

Situations in which self-organised collective action can be sustained are marked by certain key features (Ostrom, 2000). Often they rely on the emergence of a leader or entrepreneur who provides improved ways to achieve joint outcomes. Importantly, users of the collective resource (eg, a non-excludable but rivalrous social service) determine their own membership, which assists in building trust and reciprocity. Membership can be distinguished along the lines of communities with shared values and interests. In some cases this coincides with geographical, religious or ethnic communities. Other important ingredients for successful and sustained collective action include users of the collective resource being able to shape and enforce their own rules. And for larger collective enterprises it is also necessary to have multiple and nested layers of governance.

Successful collective action can be hard to sustain (Ostrom, 2000). Threats include rapid population shifts, or the emergence of corruption or other forms of opportunistic behaviour. These undermine common understandings of norms such as trust and reciprocity. They also include rapid changes in technology, and attempts by government to externally impose rules on collective endeavours. Collective action is also hard to sustain if the principles underlying collective action are not passed on to successors of the founding members.

In terms of social services, collective action to provide those services based on shared social norms represents a possible alternative to centralised government organisation of provision. This is more likely to be the case where strong social or cultural norms facilitate trust and reciprocity. Mission-oriented NFPs, and religious, iwi and other ethnic groups are possible candidates. However, when dealing with such groups it is important to recognise that their members are likely to be motivated by more than just self-interest. As Ostrom (2000) puts it:

It is possible that past policy initiatives to encourage collective action that were based primarily on externally changing payoff structures for rational egoists may have been misdirected – and perhaps even crowded out the formation of social norms that might have enhanced cooperative behaviour in their own way. (p. 154)

Conversely:

Increasing the authority of individuals to devise their own rules may well result in processes that allow social norms to evolve and thereby increase the probability of individuals better solving collective action problems. (p. 154)

## F.4 Purchasing social services – *how to buy*

Section F.3 addressed the *make or buy* question. This section considers the factors affecting how best to purchase social services from non-government providers, assuming government has decided to purchase externally. That is, the *how to buy* question.

This section highlights the optimal design of incentives in purchase contracts, the design of tenders when both cost and quality are of concern, when negotiation is to be preferred to competitive tendering (including relational contracting), and the optimal delegation of purchase decisions.

### Optimal incentive power in purchase contracts

Section F.3 emphasised contractual incompleteness as an important driver of the *make or buy* decision. In the context of purchase, agents have choices as to the optimal extent of such incompleteness. This affects their choice of purchase incentive power (Bajari & Tadelis, 2001). Under this approach, purchase is regarded mainly as a problem of adaptation in the face of changed circumstances rather than one of bidder screening (as in the simple purchase model under adverse selection discussed in Dixit (2002)). Screening at the start of the contract remains important, but with mechanisms such as competitive bidding and bidder reputation used to alleviate adverse selection problems. Instead the focus is on how the specificity of project designs can lead to future re-negotiation as circumstances and needs change.

Under Bajari and Tadelis' approach, a purchaser provides a seller with a product design. A more complete design reduces the prospect of future contract renegotiation, but is costly. The provider can exert non-contractible, cost-reducing effort. If re-negotiation occurs, the provider has private information about the cost of changes to the original design. The authors contrast fixed-price purchase contracts with cost-plus contracts. The former give strong incentives for cost reduction as well as greater risk to a provider, since the provider keeps all gains from cost reductions. By contrast, cost-plus contracts offer weak incentives for cost reductions and lower risk to a provider, since they are fully reimbursed for their costs, whatever those costs may be.<sup>42</sup>

Simple (ie, cheap to specify) projects in Bajari and Tadelis' model are optimally purchased under fixed-price contracts and with specific designs. Conversely more complex projects are optimally purchased using cost-plus contracts and less-specific designs (giving greater prospect of future design changes and re-negotiations). This stems from a trade-off between providing performance incentives at the time of contracting, and avoiding transaction costs due to subsequent renegotiation as circumstances change.

Fixed-price contracts provide strong incentives to reduce costs, but they suffer the risk of costly renegotiation when used with specific project designs. This is because the provider faces greater financial risk under fixed-price contracts when circumstances change, but project specification is fixed and cannot be adapted to those new circumstances. This risk is less important, however, when projects are simple, since it is less likely that specifications will need to be changed in response to changing circumstances. Conversely, cost-plus contracts offer weak incentives for cost reduction, but are better able to be adapted to changing project needs, and so are better suited to complex projects.<sup>43</sup>

Social services are often bundled, and their particular requirements depend on the circumstances of diverse clients. This suggests social services are both complex and hard to specify. Bajari and Tadelis' analysis therefore suggests that cost-plus social service delivery contracts might be more efficient than fixed-price contracts, all other things being equal.

## Tender design when price and quality matter

The need for transparency, value for money, accountability for expenditure of taxpayer funds, and avoidance of cronyism and favouritism, motivates the use of competitive tenders when governments purchase from private providers. The use of tenders can also be motivated on incentive grounds even for non-government purchasers, when providers have hidden information regarding their true supply costs (see the simple example given in the section F.3 discussion of adverse selection). Importantly, to be effective at achieving value for money, tenders rely on the availability of an ongoing pool of potential competing bidders.

In contrast, private-sector purchasers use negotiation more frequently than government purchasers instead of competitive tendering (Domberger & Rimmer, 1994). They also use relational contracting more frequently than government purchasers, recognising the need to take past provider performance into account (Kelman, 1990). While this does not necessarily achieve lowest cost, it results in better quality.

Indeed, it is well known that by inducing lowest bids, competitive tenders can result in quality shading (Manelli & Vincent, 1995).<sup>44</sup> More precisely, tenders are optimal when providers care more about quality than the purchaser, but result in quality shading otherwise. Conversely, in some circumstances, if the purchaser cares more about quality than providers, then arbitrarily selecting a provider and making them a take-it-or-leave-it offer can be optimal. In fact, when providers have private information about their product costs, it can be optimal for purchasers to deliberately understate their desired quality (Asker & Cantillon, 2005). This approach means that providers cannot differentiate themselves on quality and so must compete harder on cost.

<sup>42</sup> Indeed, cost-plus contracts also create incentives for over-spending – for example by unnecessarily “gold plating” outputs.

<sup>43</sup> Note that the predictions of Bajari and Tadelis differ to those of Spiller (2008) discussed in section F.3. Spiller predicts the use of specific, cost-plus contracts as a response to third-party opportunism, without distinguishing project complexity. Bajari and Tadelis make more refined predictions that allow for this added dimension.

<sup>44</sup> Kranton (2003) provides a complementary explanation for why competition can reduce quality. Specifically, competitive pricing can eliminate the price premium required to support the production of costly quality.

Tenders based on lowest cost are well-suited to purchasing simple goods or services with characteristics that are easily specified in advance and for which there is little risk of changing specifications post-tender. In other cases, however, alternatives to lowest cost tenders can be preferable. One such alternative is to use multi-criteria tenders that place pre-specified weights on cost as well as other desired characteristics of the good or service being purchased. In practice such tenders can be hard to implement, and are at greater risk of corruption and re-negotiation than lowest cost tenders (Saussier, Staropoli & Yvrande-Billon, 2009). They also result in quality shading when quality is hard to verify.

In principle it should be expected that the length of tendered contracts should align with the investment horizon of the providers. For example, based on hold-up considerations (see Box F.1), a provider might fear that it will not recover long-term investments if it can secure only a short-term contract in a tender. With the contract subsequently re-tendered, especially if it does not account for past performance or investments, there is a risk that any initial investment is held up.<sup>45</sup> Absent other solutions to the hold-up problem this will deter such investments, and bias providers towards shorter-term or less specific investments.

However, in a repeated tendering situation, reputation concerns of providers play a role in supporting the provision of non-contractible dimensions such as quality. In this setting, short-term contracts can produce benefits to be traded off against hold-up risks (Calzolari & Spagnolo, 2005). In particular, if quality is of concern to the purchaser but not critically so, shortening contract lengths and reducing the pool of competing providers can improve providers' incentives when past performance is taken into account. Providers who fail to deliver quality face the risk of missing out on future contracts, which induces better performance in earlier contracts.

Saussier, Staropoli and Yvrande-Billon (2009) emphasise three other important features of tenders – tendering costs, tender information structure, and tendering institutions. Tendering costs can deter potential providers from participating in tenders, and therefore affect tender competitiveness. As to information structure, disclosing non-public project information can improve bidding. However, while bids are lower on average with such disclosure, winning bids are not significantly so.<sup>46</sup> Conversely, disclosing bidder identities encourages collusion and reduces tender competitiveness. Further, new entrants in tenders are less informed than incumbents and may bid aggressively. This makes them more susceptible to losses and so to failure. Information release may also encourage bidders to participate just to obtain information.

Institutional capabilities such as public-sector tendering expertise affect the quality of contract specification and tender processes, in turn affecting bidding costs and re-negotiation risks. Where public bodies cannot make credible commitments, this deters good quality bidders or increases their bid prices. And if enforcement institutions (eg, regulatory bodies) are weak, this can result in greater re-negotiation and opportunism by the non-government partner. This is because weak enforcement institutions are less likely to hold bidders to previously agreed contracts, which induces those bidders to better those contracts through re-negotiation.

In the context of tenders for social services contracts potentially involving both NFP and FP bidders, the economics literature offers little specific guidance on optimal tender design. However, potentially important considerations can be identified based on the comparison of NFP and FP features in section F.3. Specifically, NFPs can be predicted to have a possible cost advantage relative to FPs when bidding in tenders, and also a possible quality advantage. They also have a possible reputational advantage when purchasers are risk-averse, but disadvantages in terms of risk-bearing.

The cost advantage flows from NFPs' ability to use alignment of mission orientation with employees to reduce the monetary rewards they must offer to secure them. Indeed, NFPs often enjoy voluntary labour, with mission alignment singly important in attracting staff. On the face of it NFPs should therefore have lower wage costs for each worker than FPs. However, countering this is the selection issue highlighted in

<sup>45</sup> One possible solution might be to reimburse the provider's investment costs upfront. However, this could weaken the provider's performance incentives, since they would then have less at stake when providing their services. Also, reimbursing investment costs could induce the provider to over-invest. This means that any upfront reimbursement of investment costs is likely to be problematic, even if it resolves hold-up risks.

<sup>46</sup> One explanation is that disclosure means the value of a bidder's private information is reduced, since all bidders potentially have access to comparable information.

section F.3. Specifically, FPs can offer higher pay than NFPs, and therefore can cherry-pick the most able employees. This means the productivity of each employee in FPs might be higher than in NFPs, countering the apparent labour cost advantage of the latter.<sup>47</sup>

As to quality, and also as discussed in section F.3, NFPs represent a commitment device to offering relatively low-powered incentives and therefore lower quality shading. Mission alignment with workers also supports non-monetary incentives for quality provision. Once again this advantage relative to FPs must be weighed against the ability of the latter to attract more able workers, as well as reputational concerns of FPs supporting a commitment to quality. However this is less likely to reverse the quality advantages of NFPs relative to FPs. So in tenders for which hidden quality is particularly an issue, NFPs might retain a bidding advantage (ie, if their bid cost matches or is even higher than that of an FP rival, their expected quality provision could be higher).

NFPs potentially possess a reputational advantage relative to FPs. This is because they are often regarded as well-meaning and socially beneficial, as opposed to self-interested and potentially at odds with social preferences. This means that risk-averse providers might take comfort from such a reputation when contracting with NFPs, since it reduces the risk of political embarrassment as to provider choice.

NFPs have restricted access to financial capital, relative to FPs. This may exclude most or all NFPs from tendering for capital-intensive projects. For projects with more modest capital requirements, FPs will generally face a lower cost of capital, with a corresponding cost advantage over NFPs.

NFPs should be assumed to be more risk averse than FPs (section F.3). Tenders are commonly for provision at a fixed price and, as in the Bajari and Tadelis (2001) discussion above, this imposes high levels of risk as well as incentives on the provider (particularly when specifications of the relevant good or service can change post-tender). This places NFPs at a bidding disadvantage relative to FPs, since relative risk-aversion considerations favour them offering a higher bid than FPs so as to reduce their exposure to risk. A possible solution to this is the use of collaborative bidding (ie, with other NFPs or FPs) to share and reduce risk, though this reduces the competitiveness of the tender by reducing the number of effective bidders.

## Negotiation and relational contracting

The above discussion of optimal incentive power in purchase contracts highlights that fixed-price purchase with high-powered incentives is not always to be preferred. Indeed, low-powered, cost-plus contracts are favoured when purchasing complex goods or services. Since such contracts do not specify a cost, they are not conducive to being purchased competitively via cost-based tenders. While in principle cost-plus contracts could be tendered on non-cost dimensions (such as quality), this is harder to specify and verify, and so diminish any advantages of competitive tendering. This favours consideration of alternative purchase methods such as negotiation and relational contracting.

Non-competitive negotiation between a purchaser and provider can be preferred when the purchaser relies on the experience and expertise of the providers in designing or specifying the good or service to be purchased (Saussier, Staropoli & Yvrande-Billon, 2009). Indeed, where such design or specification involves the generation of intellectual property, negotiation can be used as a means to reduce the risk of that intellectual property being disclosed to other parties.

Another reason for favouring negotiation arises when purchase is repeated, in which case non-contractible dimensions such as quality can be addressed through reputation concerns and relational contracting (Calzolari & Spagnolo, 2005). As discussed above, a purchaser can provide incentives for a provider to maintain supply quality through the threat of being excluded from future purchase decisions. When quality is particularly important to the buyer, this can favour exclusive, relational contracting with just one supplier.

An oft-cited example of the more relational approach suggested by Calzolari and Spagnolo is the *keiretsu* system used in Japanese industry, for example, by Toyota (MacDuffie & Helper, 2005). Under this approach, purchasers maintain relationships with a small set of suppliers, combining information sharing, close

<sup>47</sup> It is also possible that mission orientation provides such strong incentives that well-paid non-mission oriented employees are no more productive than lower-paid, mission-oriented employees. In that case, the cost advantage of NFPs in tenders would be reinforced rather than undermined by lower monetary incentives in NFPs.

monitoring and limited competition (Williamson, 2002). This is to be contrasted, for example, with the traditionally more hard-nosed approach of US car manufacturers based on arm's length relationships, competitive bidding and lowest cost, although there is increasing convergence between these two models (MacDuffie & Helper, 2005).

As identified above, when providers must make long-term investments there is a trade-off between hold-up risks and inducing quality incentives when choosing contract length in tenders. In principal such fixed investment costs might be better addressed through long-term contracts, which could be either tendered or negotiated, with negotiated contracts favoured when providers have diverse contractual requirements (ie, contracts are too non-standardised to be effectively tendered). However, Asker and Cantillon (2005) model optimal purchase when providers have private information about not just their per-unit quality costs, but also their fixed costs of quality. They show that multi-criteria tenders can remain superior to negotiation in this case.

Negotiation instead of tendering carries obvious costs and risks. A lack of competition means negotiated contracts may not be least cost. Further, more relational contracting carries the risk of cronyism and favouritism, neither of which are consistent with public sector accountability, and which raise political risks. In choosing between the two modes of purchase, these costs must be weighed against the relative merits of each mode.

## Competition for the market versus competition in the market

How government purchases services from non-government providers affects the nature, extent and timing of competition in the delivery of social services. In particular, if government offers an exclusive contract to one provider, then this limits the competition (as well as opportunities for collaboration) faced by that provider over the life of the contract. Conversely, if government simultaneously offers multiple contracts, this supports greater competition and/or collaboration between providers.

If a single contract is offered by way of competitive tender, this is an example of *competition for the market*. While only a single provider can win the tender, providers compete to offer the best possible terms to secure the exclusive contract (unless they collude or collaborate in the tender). The competitive pressure facing the winning provider changes over the life of the contract, with pressure for performance diminishing once the contract has been awarded, but growing as the contract nears maturity and becomes available for re-tendering.

An advantage of competition for the market is that competition for the contract can be intense, offering reduced purchase costs and possibly more innovative ideas for service delivery. It also offers the successful provider a relatively low-risk environment within which to deliver their services, which can support experimentation. The provider also benefits from the possibility of gaining an "incumbency advantage", meaning that by becoming better-known by the purchaser they might improve their chance of contract renewal upon its termination. This assists them in recovering the costs of multi-period investments, and so can improve service provision. Incumbency advantage also rewards good contractual performance, and so provides additional performance incentives beyond those explicitly provided for in the contract.

A disadvantage of tendering single contracts is the difficulty in ensuring the survival of, or entry by, potential bidders in future tenders. If an exclusive contract is offered for a sufficiently long period, it may be impossible for other providers to remain active in the relevant service area. Further, any incumbency advantage enjoyed by a winning provider makes it harder for new providers to win subsequently tendered contracts. This means that the desirability of competition for the market depends on the balance between reduced delivery costs and any improved incentives offered by reduced competition and incumbency advantage on the one hand, and the potential costs of reducing competition in future tenders on the other.

When multiple contracts are simultaneously tendered, this is a form of *competition in the market*.<sup>48</sup> Having multiple providers offers potential advantages in terms of greater diversity of service provision. This implies

---

<sup>48</sup> Conventionally, competition in the market refers to a situation in which multiple providers compete on price, variety and quality on a continuous basis. When governments tender multiple contracts simultaneously, however, this involves bidders competing to secure contracts with generally fixed prices and quality. Competition over the life of those contracts is then restricted to other dimensions such as the number of clients the providers can attract. This represents a hybrid form of competition in the market.

diversified risk to the purchaser of poor service delivery. It also means that there is reduced incumbency advantage, and a greater chance that future contract tenders will remain competitive (due to more potential providers remaining active in the relevant service area). Also, since each contract is smaller than if a single contract had been tendered, this implies that the setup costs for entrants will also be lower, improving the likelihood that new providers will participate in future tenders.

A possible disadvantage of tendering multiple contracts is that the number of bidders for each contract might be lower than if a single contract was tendered.<sup>49</sup> This is particularly the case if there are fixed costs of tender participation, which are less likely to be recoverable with smaller contracts being offered. This means the purchaser might secure less-favourable prices for each contract.<sup>50</sup> Additionally, tendering multiple small contracts might be inefficient compared with tendering a single, large contract. This is particularly if there are economies of scale in service provision.

Tendering smaller contracts has the disadvantage that the possible gains from winning the contract might be insufficient to justify any fixed costs of tender participation. The desirability of competition in the market therefore trades off the possible advantages in terms of greater service diversity and survival of possible competitors for future contracts on the one hand, against possible losses in competitive pricing and economies of scale on the other.

## Optimal delegation of purchase decisions

Assuming social services are to be purchased instead of provided by government, a related question is where best to locate the decision to make or buy within government. The decision can be centralised in the hands of politicians (principals) or delegated to subordinates (agents).

Delegating purchase decisions shares considerations relevant to delegating monetary policy decisions to central banks. In particular, developed countries have increasingly delegated monetary policy decisions to independent central banks. This is to insulate monetary policy decisions from short-term political pressures, and by so doing resolve the monetary policy equivalent of hold-up, known as policy time-inconsistency (Dincer & Eichengreen, 2014).

Similar considerations apply in the context of purchase. Delegating purchase decisions to bureaucrats rather than politicians insulates those decisions to some degree from political pressures. Indeed, purchase decisions can become based more on technical (eg, economic) than political grounds when purchase decisions are delegated to parties with preferences and incentives located between those of politicians and non-government providers. Further, delegating purchase decisions to such parties can improve providers' performance incentives by giving them a greater sense of control (Dessein, 2002; Aghion & Tirole, 1997). Politicians therefore face a trade-off between improving performance incentives through delegation, and the loss of control that delegation entails.

Bennedsen and Shultz (2011) consider strategic delegation under incomplete contracting, in which incentives are given only indirectly via post-contract renegotiation. They assume that government faces a cost-quality trade-off when choosing between contracting with private or non-government providers, with delegation having two important effects – incentive and bargaining.

- *Incentive effect* – the principal can influence a service-provider's incentives by delegating the future contract re-negotiation to an agent that contracts with the provider and which has different preferences to those of the principal (eg, regarding cost reductions). Due to assumed contractual incompleteness, the contract will be re-negotiated after it has been agreed. However, the provider's performance incentives are less affected when anticipating this re-negotiation if the provider anticipates dealing with

<sup>49</sup> In fact the number of bidders in any given tender will reflect not only the number and type of contracts on offer, but also the prospect of winning a contract in the tender. If the number of contracts on offer increases, this suggests a greater chance of winning the tender for a given number of bidders. However, additional bidders might be induced to compete for these contracts, supporting the competitiveness of each tender, for example if setup costs for smaller contracts are more bearable.

<sup>50</sup> As above, however, this might relieve possible quality shading issues.

an agent that has preferences more aligned than the principal's preferences with its own. Delegation therefore substitutes to some degree for an explicit incentive contract.<sup>51</sup>

- *Bargaining effect* – an agent's preferences for cost and quality affect their preference for purchase, given differences in cost-quality trade-offs under government provision and external purchase. The principal exploits this by delegating to an agent that prefers not to outsource, forcing private firms to price more competitively. This makes delegation a credible commitment to tougher bargaining. Who to delegate to becomes the important political choice.

Politicians face two sets of decisions – whether or not to outsource, and to whom to delegate re-negotiation. Under *partial delegation* the principal decides on whether or not to outsource, while re-negotiation with the provider is delegated to the agent. This is shown to be the most efficient approach. Conversely, under *arm's length delegation*, both decisions are delegated to an agent. Efficiency impacts in this case are mixed.

Bennedsen and Shultz conclude that partial delegation results in outsourcing being optimal under a wider set of parameters than arm's length delegation.

## F.5 Innovation in social services

### Theories of innovation

Economist Joseph Schumpeter promoted the idea that innovation is an important way in which firms compete (OECD, 2005). He described innovation as a process of *creative destruction* in which organisations adopting new technologies displace firms using older technologies. Innovation can be either incremental or radical/disruptive. Incremental innovation gives rise to continuous, small changes that are relatively low risk, and which can be achieved in the short term. Conversely, radical innovation – Schumpeter's main focus – leads to fundamental changes in products, markets, production processes, input supplies or market structures. Radical innovation is inherently higher risk, and can take a long time until fruition.

Systems innovation theory extends Schumpeter's ideas beyond competing firms. It instead highlights systemic interactions that support innovation (Lundvall, 1992). In particular, innovation is not just an instrument of competition. Rather, it often arises from coordination and cooperation among multiple organisations. Such organisations include private research organisations, but also government research bodies and universities.

### Elements of innovation – adaptation, experimentation and diffusion

#### Adaptation

Both incremental and disruptive innovation can be decomposed into further elements. One is adaptation, meaning the way in which goods or services, or their means of production, are tailored to particular market niches. Such tailoring might be only incremental in nature. However, it could also be a disruptive form of innovation, such as when the preferences of the relevant niches have not previously been served.

Where information about the preferences of these niches can be easily discovered or held centrally, then hierarchical organisations can effectively provide such adaptation. However, such preference information can be hard to obtain. This could be, for example, because social services clients are reluctant to share it with large, impersonal hierarchies. In this case, preference information is more likely to be held in smaller, more local organisations, placing them in a better position to provide adaptation. A cost of this information being held at smaller scale or more locally is that adaptations are limited in their scale or reach.

#### Experimentation

Another element of innovation is experimentation. This includes trialling one or more changes in existing products or services, client groups, delivery methods or organisational forms to see if these changes improve outcomes. Experimentation involves trial and error, and necessarily implies departures from the tried and true, so it is naturally risky. Further, it requires that the parties organising the experiments have the capacity to organise, implement and monitor them.

<sup>51</sup> Similar incentive effects are discussed in section F.3 in relation to Aghion and Tirole (1997) and Dessein (2002).

The incentives facing experimenters are critical. For example, if the gains from experimentation are not captured by the parties taking the associated risks, then this deters experimentation. This is particularly the case when social services are purchased via competitive tender, for example. If potential providers present new ideas to purchasers, and those ideas then find their way into tender specifications, then the party presenting the new ideas is forced to compete with other providers to implement them. If that party does not win the tender, then they cannot recover the costs of developing their ideas. Conversely, if the tendered contract is awarded to more than one party, the provider that developed the new ideas might recover too little of the benefits to reward those new ideas. Either way, that party's investment in new ideas is held up (Box F.1), which deters investment in new ideas.

## Diffusion

Another element of innovation is diffusion. This is the process by which new ideas are implemented, and therefore scaled up (eg, relative to a localised, experiment). As just discussed, diffusion can occur via the purchase process itself, when the intellectual property in new ideas are not reserved to the innovator. Conversely, diffusion can occur via innovators scaling up their own production. This requires the innovator to have the resources and skills to do so, which might constrain the rewards to innovation by limiting the possible scope of diffusion.

An alternative to such internal or *organic* diffusion is to rely on the resources and skills of other parties to implement the new ideas. This could involve the innovator selling or licensing their new ideas to third parties. Ownership of innovators might even be acquired by other parties who then implement their new ideas at scale. Either way, incentives for creating new ideas are preserved, since they are compensated by those other parties using the new ideas.

In other cases diffusion is achieved by new ideas being *open source* rather than proprietary as in the preceding examples. This means the ideas are freely developed for use by others without direct compensation to the developers. Open source innovation is used principally in the IT sector, where open source software developers are motivated by intrinsic pleasure as well as through enhanced future job prospects (Lerner & Tirole, 2005). Future job prospects are enhanced particularly where the developers' efforts are highly visible to relevant audiences and strongly linked to performance. This is also the case when performance is informative about the developers' talents.

The diffusion of innovations also relies upon organisations' "absorptive capacity" in relation to new technologies. Technology diffusion, internationally, is driven by a small number of highly innovative countries whose technology diffuses to other countries via their openness to imports (Keller, 2004). However, technology adoption requires that the importing country also has the requisite local skills and resources. Further, Keller comments that "technological knowledge spill-overs appear to be resulting from a deliberate commitment to learning and matching international performance standards through ongoing interaction with foreigners" (p. 779). This emphasises the importance of looking to models of social services delivery in other countries and having strong local technology absorptive capabilities to benefit locally from international innovations.

## Organisational features affecting disruptive innovation

Yu and Hang (2009) survey organisational features affecting disruptive innovation. They highlight how human resources, and organisational culture and structure, affect the ability of organisations to develop or respond to disruptive innovations.

### Human resources

Human resource aspects are particularly important in this regard. Innovation requires risk-taking employees with the ability to think beyond established practices and rules. Often disruptive innovation arises from new entrants challenging incumbent organisations, or from frustrated innovators leaving such organisations to find the freedom to develop their new ideas. Long-term incentive plans are critical for such innovators, which can create tensions within organisations that rely on short-term, egalitarian and/or formulaic incentive arrangements. This is particularly the case since disruptive innovations are hard to anticipate or specify in contracts, and can result in major improvements to organisational performance. Also, innovations are often driven by information gained from close contact with clients, which is at lower tiers in hierarchical

organisations. Middle managers in such hierarchies can face limited benefits but high personal costs from disruptive innovation, and so favour innovations that support their ongoing roles.

### **Organisational culture**

Organisational culture – including mission orientation (Box F.5) – provides an informal way of controlling and coordinating the behaviour of employees. However, this also creates cultural inertia, and means that disruptive innovation requires employees to “unlearn deeply entrenched values” (Yu & Hang, 2009, p. 9). Indeed, organisations relying on alignment of mission orientation to compete for workers and customers can resist change for fear of creating mission divergence (Besley & Ghatak, 2005). This also causes them to be conservative and resistant to disruptive innovation. Conversely, organisational culture supports innovation when it embodies elements such as entrepreneurship, risk-taking, creativity and flexibility.

### **Organisational structure**

Organisational structure can play a key role in facilitating disruptive innovations. Yu and Hang (2009) highlight how successfully innovating organisations often created autonomous units for developing and diffusing innovations. This creates space for innovators to behave in ways not suited to the rest of the organisation (eg, ignoring established practices). It also enables different performance standards to be applied – rather than using performance criteria developed for established activities to evaluate relatively novel, risky and speculative innovations.

### **Institutional features – intellectual property rights**

A key institutional feature affecting incentives to innovate is the availability of effective intellectual property rights, including the effective enforcement of such rights. Indeed, Falk (2006) finds that protection of intellectual property rights is unambiguously positive for innovation intensity across a range of empirical studies. Such rights create time-limited market power for innovators, meaning they can secure returns from their innovations – either directly or by allowing others to use their ideas. These returns then enable recovery of the costs of developing innovations. The trade-off that this creates is that innovations will not diffuse as quickly as if there were no intellectual property protections. However, without such protections there might be fewer innovations to diffuse. In other words, there is a trade-off between incentives to innovate, and incentives to diffuse innovations.

Assuming effective property rights are available to secure the new ideas of innovators, a related question is how those rights should be assigned. If governments purchase services from third parties via fixed-price competitive tenders, this assignment critically affects innovation incentives. If providers must cede their intellectual property rights in new ideas to government as part of their contract, then this caps the benefits they can secure from those ideas. However, if providers can fully retain those property rights, this leaves them with the greatest incentive to innovate.<sup>52</sup> Indeed, leaving the rights to new innovations fully with providers could induce them to bid more aggressively to secure contracts in the first place – trading off short-term losses in revenues against possible long-term gains.

### **Risk sharing**

A key feature of innovation is that it has highly uncertain outcomes (Smith, 2006). This demands the use of compensating risk-sharing mechanisms to encourage risk-averse organisations to innovate.

One such mechanism is the use of research joint ventures or other forms of collaborative research (Smith, 2006). Indeed, research collaboration is favoured for sharing both the costs and risks of innovation, to enhance market power, as well as to internalise the technology spill-overs that commonly arise with innovation (Caloghirou, Ioannides & Vonortas, 2003).

---

<sup>52</sup> This does not, however, assure that they have the resources to innovate. Possible mechanisms to balance the incentive and means to innovate include governments offering financial prizes for, or purchasing, innovations. Alternatively, government might pay for a licence to use the innovations of providers.

## Competition

Consistent with Schumpeter's theory, in practice competition is commonly found to be a key driver of innovation.<sup>53</sup> Specifically, competition has pervasive and long-term effects on economic performance by encouraging the innovative activities of firms, and by selecting more efficient activities over less efficient ones over time (Ahn, 2002). However, the relationship between competition and innovation is not simple, with competition found to have an inverted U-shaped relationship with innovation (Aghion et al., 2005). As competition increases from low levels, firms innovate to escape the effects of increased competition. Conversely, returns to innovation are not captured by innovating firms when competition is high, meaning that increases in competition beyond some level cause innovation to decline. This means that there is some intermediate level of competition at which innovation is at its maximum, trading off these two effects.

This suggests that introducing some degree of competition in the purchase of social services might be beneficial for innovation. However, the existing economics literature offers little specific guidance on how competition between NFP, FPs and government providers affects incentives to innovate, especially for services with the particular characteristics of social services.

## Ownership form and innovation

### Government

Many features of government affect its ability to innovate, and how and what forms of innovation it can sustain. These include the inability of government to commit to incentives for innovations made by government workers, the risk of third-party opportunism, and government structure and culture.

As discussed in section F.3, governments have less ability than FPs to commit not to expropriate the returns from the innovations of employees (Hart, Shleifer & Vishny, 1997). This places governments at a disadvantage relative to FPs in relation to innovation.

Further, because innovations are highly uncertain, incentive contracts for innovations are particularly hard to specify in advance. Such contracts therefore suffer from potentially severe incompleteness issues, meaning it is hard to use just contracts to provide long-term incentives for innovation. Since governments cannot offer their employees ownership-based incentives, such as shares or share options, this means governments lack an important alternative to contracts for providing the incentives needed for innovation. Indeed, share-based performance rewards are associated with more successful or higher-value innovation (Lerner & Wulf, 2007).

Also, as discussed in section F.3, the risk of third-party opportunism (eg, attacks by political rivals) causes governments to be highly risk-averse. As a consequence they are more specific in their contracts, and more rigid in the implementation of those contracts (Moszoro & Spiller, 2012; Moszoro, Spiller & Stolorz, 2014). This naturally limits experimentation, since it is inherently risky, with perceived failures of experiments opening up governments to criticism.

Indeed, accountability and transparency requirements induce governments to resort to the relative safety of "best practice" in purchase and provision activities. By definition, best practice is retrospective, since it relies on the current state of the art, rather than the emerging state of the art (ie, innovation). In addition, requirements to achieve perceived value-for-money in purchase can be taken to the extreme by zealous government employees, including demanding all intellectual property rights from providers of purchased goods or services.

Third-party opportunism also induces governments to hire employees that are safe and rule-abiding, rather than risk-taking and not rule-bound. Combined with governments' use of typically low-powered incentives, this means governments are unlikely to recruit employees with the characteristics most required for innovation. Indeed, the need for centralised control of risky government activities limits the degree to which government employees can be given the autonomy to innovate. This further impedes government's ability to recruit employees with the skills and characteristics needed for innovation.

<sup>53</sup> Contrast that with the discussion in section F.3, where it is noted that competition need not unambiguously improve performance incentives (as opposed to innovation).

Government culture and structure also inhibits innovation. The public service ethos is predominantly one of conservatism and risk-limitation rather than innovativeness and risk-taking. Further, this culture is deeply ingrained and so is slow to change in the face of disruptive innovations.

Government's dominance in purchasing social services means that it has both the incentive and means to limit disruptive innovations. Government employees can favour less-disruptive innovations if this protects their tenure and status. They can influence contract specifications to ensure this is achieved. Likewise, government has an interest in maintaining competition in tenders. It can also achieve this by specifying purchase contracts so as to limit provider innovations that threaten the viability of other providers.

All of these features suggest that governments are not naturally able to support innovation beyond relatively safe and incremental innovations, all other things being equal.<sup>54</sup> Indeed, in practice, governments can be relatively slow to adopt new technologies (Rose & Joskow, 1990).

### **Not-for-profits**

Like governments, NFPs have features that limit their ability to innovate. They are less exposed to third-party opportunism than governments, and also less likely to expropriate the value of employees' innovations. However, they share limitations in terms of being able to offer long-term incentives for innovations, and in terms of risk aversion.

Unlike FPs, NFPs are unable to raise new capital by issuing ownership claims to investors. This means they face capital constraints, which increase their risk aversion.<sup>55</sup> As for governments, this naturally hinders experimentation. A lack of capital also impedes the ability of NFPs to invest in innovative activities, or to diffuse new innovations through organic growth. In each case NFPs rely on cross-organisational diffusion instead.

NFPs cannot issue shares or share options to their employees. Because of their constraints on distributing profits or reserves, they also find it difficult to offer employees performance-related pay. This means that, like governments, NFPs cannot provide long-term, high-powered incentives for innovation to their employees.

Indeed, to a greater extent than governments, NFPs rely on the mission orientation of their employees to encourage performance. In principle, this orientation could provide strong incentives for innovation. However, this must be weighed against the rigidity associated with organisations having strong cultural values. The mission orientation of NFPs could even support open-source innovation, with the returns from an NFP's innovation shared with other parties in the expectation that its own aims will be furthered. However, this must be weighed against the need for NFPs to continue to secure supply contracts from government to remain financially viable. Sharing innovations with rivals risks NFPs becoming less competitive when bidding for government contracts.

Given these features, NFPs are likely to also face obstacles in creating or implementing disruptive innovations, all other things being equal. As for government, but for different reasons, they are likely to be better able to support incremental innovations.

### **For-profit firms**

FPs are most able to engage in innovation of the three organisational forms considered. They can issue shares to investors and other organisations. This enables them to raise the external capital required to fund innovation, including the organic diffusion of innovations. It also enables FPs to more easily participate in joint ventures for sharing the risks and costs of innovative activities.

Since they are not subject to the same political scrutiny as governments, FPs are less exposed to risks of third-party opportunism. This means they retain greater flexibility in how they are structured, including a greater ability to decentralise control of innovative activities to autonomous units. They also rely less than governments on tried and true systems, meaning they are more flexible and can adopt new technologies

<sup>54</sup> Major government-funded aerospace and military innovations represent important exceptions. However, they are of limited relevance to a discussion of innovation in social services. In any case, even such innovations have often relied on innovation by FPs (eg, the B-2 Stealth bomber development by Northrop, Boeing, Vaught and General Electric (Cremer, Garicano & Prat, 2004)).

<sup>55</sup> To some degree these constraints might be relaxed by using financing innovations such as crowd funding. Arguably NFPs have an advantage relative to FPs and governments in accessing such funding, given their ability to rely on mission orientation alignment to attract funders.

more easily. Additionally, the reduced exposure of FPs to third-party opportunism as well as greater access to capital mean they are better able than governments to bear the risks associated with experimentation.

FPs lack the mission orientation of governments and NFPs. This means they are more likely to favour proprietary over open-source innovations. It also means they rely more heavily on high-powered incentives to induce innovation by employees. This means, however, they can also cherry-pick employees with the skills and characteristics needed for innovation. And because they can offer share-based compensation to employees, this credibly provides the long-term incentives for innovation that governments and NFPs cannot provide.

These features suggest that both incremental and disruptive innovation are more likely to occur in FPs than in either governments or NFPs, all other things being equal. Indeed, while public R&D is substantial in many countries, most R&D is privately funded (Keller, 2004).

## F.6 Choice and market power, collaboration and collusion

This section briefly discusses two other topics relevant to the economics of social services. The first addresses increasing consumer choice and its impact on provider market power. The second is collaboration between providers, which might also be termed collusion in certain circumstances.

### Consumer choice and market power

A feature of the traditional delivery of social service is the limited role played by consumer choice (section F.2). Instead the quantity, quality and variety of social services are centrally determined by government, reflecting the preferences of government and providers more than those of social services clients.

Increasing consumer choice is desirable to better tailor service provision with specific client preferences. In principle this can be brought about by opening up service provision to competing providers. However, a consequence of offering greater consumer choice is that the provision of services becomes less uniform, with providers differentiating on non-price features. In general this should be expected to afford a degree of market power to providers, since they compete less directly when faced by differentiated rivals. Greater choice can be associated therefore with increased competition, but among more differentiated rivals.<sup>56</sup>

### Search costs and switching costs

A complication of offering greater provider choice to social services clients is that they then face both search and switching costs. Search costs arise when consumers must expend resources to identify their most suited providers. Switching costs result from such search costs, since a client of one provider must invest resources in locating a different provider should they wish to change. However, switching costs arise more generally. For example, establishing client–provider relationships can impose fixed costs on clients, such as the costs of providing case histories to providers. Search and switching costs provide other reasons why greater choice can result in greater provider market power.

Since searching is costly, social services clients will rationally incur those costs only if they expect to gain by doing so. This means they can remain less than fully informed about provider choices even after researching providers. This can be in the interests of providers, since it makes it more difficult for their clients to switch providers.<sup>57</sup> Clients do not necessarily suffer as a consequence of such lack of information. This is because the associated rewards from client lock-in can induce providers to compete harder and offer clients favourable terms to secure their custom in the first place (Farrell & Klemperer, 2007).<sup>58</sup>

Mechanisms to relieve search costs include producers providing warranties, and advertising (Tirole, 1988). High-quality producers have an incentive to provide warranties, as these signal the producer's quality. Poor-quality producers do not share this incentive, but their failure to offer such warranties effectively signals their

<sup>56</sup> Another issue raised by differentiation is that of collusion between social services clients and providers. For example, a client not willingly seeking employment might self-select a provider that is less likely to actively provide them with employment opportunities. Prendergast (2007) shows that it can be optimal in some cases to have empathetic providers in certain roles (eg, teachers, social workers), but antipathetic ones in others (eg, police). Ensuring such selection by provider type (ie, empathetic, antipathetic) should alleviate risks of client–provider collusion.

<sup>57</sup> Indeed, it can even be a strategy of business to confuse clients about their choices – so-called *confusopoly*.

<sup>58</sup> Loyalty rewards are another example. The risk of losing loyalty rewards when changing suppliers can cause customers to become locked in. However, their provision provides clients with benefits they might not enjoy otherwise.

low quality to consumers anyway. However, only limited warranties can be offered when there are consumer moral hazard issues, such as consumers mistreating goods in the knowledge they are covered by a warranty. Adverse selection issues also restrict feasible warranty coverage, since consumers who are more likely to need the warranty are more likely to select producers that offer them. In the social services setting both moral hazard and adverse selection issues are likely to be so severe that warranties from voluntary producers are infeasible. However, New Zealand's statutory product and service guarantees under the Consumer Guarantees Act 1993 mean that providers remain confronted with these issues.

Advertising can lower consumer search costs by providing information on the existence and cost of services. However, since advertising is costly, this information will only be provided imperfectly. Further, advertising costs have both direct and strategic effects (Tirole, 1988). They directly affect the profitability of providers. But advertising costs also allow providers to remain differentiated in informational terms, since incomplete advertising means consumers cannot easily compare all providers. This means those advertising can charge higher prices in a market setting, and explains why some professions favour advertising bans.

### **Provider choice when quality can only be learned by experience**

Providers can suffer when clients are uninformed about their characteristic (such as quality), and those characteristics can only be ascertained when clients experience their services.

This can give rise to the famous "lemons" problem (Akerlof, 1970). If quality is unknown to clients when they make their purchase decision, then providers cannot charge for quality provision in a market environment. Absent solutions to this, only providers of low quality can serve the market, resulting in low-quality service provision. Common solutions to the "lemons" problem include governmental or industry quality controls, minimum quality standards, occupational licensing, certification, and safety regulations. Both certification and occupational licensing benefit consumers who highly value quality, but do so at the expense of consumers who do not (Tirole, 1988).

Mandatory information disclosure is another possible solution for informing clients about the quality of providers (such as requiring providers to display their qualifications). However, high-quality providers already have an incentive to make such disclosures, and low-quality providers can be discerned by their non-disclosure of such information (Bolton & Dewatripont, 2005). Mandatory disclosure rules therefore can be superfluous, and even encourage excessive disclosure.

However, mandatory disclosure rules are useful when providers are unsure of their quality and must make costly investments to determine it (eg, in determining their compliance with accreditation standards). Providers can have incentives to make such investments even when they are not socially beneficial, because they enjoy private gains from learning their type. Forcing all providers to reveal such information removes this incentive.

Repeated transactions with a provider offer a way for a client to learn the quality of their provider. It also means that once they have a provider, a client will not regard alternative providers as equivalent since they have no experience of those other providers. This provides incumbent providers with an advantage over rivals, including those with potentially higher quality (Tirole, 1988).

The challenge for providers is therefore to induce clients to try their services. Common methods in market situations include offering introductory discounts or free samples. They also include other conspicuous initial expenditures (such as advertising) which represent investments in the provider's brand. These investments signal that the provider intends to remain active in the long term, and help it to develop a reputation for quality. This approach is effective when a provider risks not having repeat purchases if discovered to be of low quality (Tirole, 1988).

### **Consequences of providers enjoying market power**

To the extent that greater choice results in increased provider market power, this might be reflected in provision costs higher than those in a less differentiated market. For the reasons discussed above, however, this is not necessarily a disadvantage. Affording providers with some degree of market power is necessary when they incur fixed costs of delivery that cannot be recovered when prices are competitive. Also, a reduced risk of losing customers through switching to substitutable providers preserves incentives for

providers to make customer-specific investments.<sup>59</sup> Such investments should enhance the value of social services clients having greater choice over providers.

Intense competition between undifferentiated providers for differentiated customers could induce cream-skimming behaviour. This is because competitive producers receive low prices, and so have incentive to seek out the lowest-cost clients to serve. This risks leaving higher-cost clients unserved. Affording a degree of market power to providers can preserve their profitability even for higher-cost clients, and so support greater market coverage.

## Collaboration and collusion

Collaboration between providers can be beneficial in terms of enabling the exploitation of scale economies, such as in shared IT infrastructure. It also facilitates risk-sharing, which enables risk-averse providers to better bear strong incentives for performance (section F.3). This is a possible benefit when purchase is via fixed-price tenders, which involves such strong incentives (section F.4). And intelligent bundling of activities can also give rise to complementary incentives, for example, where efforts on one activity reduces the cost of efforts on another. Indeed, this is just one way in which collaboration might give rise to economies of scope. Collaboration can also bring benefits in terms of sharing the costs and risks of innovation, and internalise spill-overs from innovation (section F.5).

In some contexts, however, collaboration can instead be thought of as collusion. Specifically, collaboration between bidders reduces the effective competitiveness of tenders (section F.4). Indeed, collusion in the context of tenders can also involve bid-rigging, which is commonly illegal. Further, collaborating providers might collude so as to reduce effort levels when subjected to relative performance evaluation, reducing any payoffs for the purchaser.

As when market power is induced by greater customer choice, there can be benefits to government even when provider collaboration takes on elements of collusion. For example, reducing tender competitiveness can leave extra returns to providers that support investments and reduce quality shading. This means that government faces a tension when determining the optimal competitiveness of a tender. On the one hand, it gains from greater competition in the form of lower purchase costs. On the other hand, it can benefit from reduced quality shading when tender competition is reduced. More generally, providers agreeing to reduce efforts on relatively unproductive tasks might induce greater efforts on other more productive but less-observable ones. This could potentially improve outcomes for the government as purchaser.

As for the tension between choice and market power, finding the right balance between these tensions will hinge on the specifics of the relevant social service.

## References

- Acemoglu, D., Kremer, M., & Mian, A. (2007). Incentives in markets, firms and governments. *Journal of Law, Economics & Organization*, 24(2), 273-306.
- Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted U relationship, *The Quarterly Journal of Economics*, May, 701–28.
- Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations, *Journal of Political Economy*, 105(1), 1-29.
- Ahn, S. (2002). *Competition, innovation and productivity growth: A review of theory and evidence*. OECD Economics Department Working Papers No. 317. Retrieved 28 January 2015 from <https://ideas.repec.org/p/oec/ecoaaa/317-en.html>
- Akerlof, G. (1970). The Market for "lemons": Quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, 84(3), August, 488–500.

<sup>59</sup> For this reason it can be desirable to restrict client switching after they have made their initial choice of provider.

- Alchian, A., & Demsetz, H. (1972). Production, information costs, and economic organization, *American Economic Review*, 62, 777–95.
- Asker, J., & Cantillon, E. (2005). *Optimal procurement when both price and quality matter*. CEPR Discussion Paper No. 5276. Retrieved 16 January 2015 from <http://ssrn.com/abstract=851190>
- Bacchiega, A., & Borzaga, C. (2001). Social enterprises as incentive structures, in Borzaga, C., & Defourny, J. (eds.), *The Emergence of Social Enterprise*. London and New York: Routledge.
- Baker, G., Gibbons, R., & Murphy, K. (2002). Relational contracts and the theory of the firm, *Quarterly Journal of Economics*, 117(1), 39–84.
- Bajari, P., & Tadelis, S. (2001). Incentives versus transaction costs: a theory of procurement contracts. *RAND Journal of Economics*, 32(3), Autumn, 387–407.
- Barr, N. (2012). *Economics of the welfare state*. 5th ed. Oxford: Oxford University Press.
- Beiner, S., Schmid M., & Wazenried, G. (2011). Product market competition, managerial incentives and firm valuation. *European Financial Management*, 17(2), 331–66.
- Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation, *Review of Economic Studies*, 70, 489–520.
- Bénabou, R., & Tirole, J. (2013). Bonus culture: Competitive pay, screening, and multitasking, Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit, No. 7321. Retrieved 26 January 2015 from <http://hdl.handle.net/10419/71723>.
- Bennedsen, M., & Schultz, C. (2011). Arm's length delegation of public services, *Journal of Public Economics*, 95, 543–52.
- Bennett, J., Iossa, E., & Legrenzi, G. (2003). The role of commercial non-profit organizations in the provision of public services, *Oxford Review of Economic Policy*, 19(2), 335–47.
- Besley, T., & Ghatak, M. (2001). Government versus private ownership of public goods, *Quarterly Journal of Economics*, 116(4), November, 1343–372.
- Besley, T., & Ghatak, M. (2003). Incentives, choice and accountability in the provision of public services, *Oxford Review of Economic Policy*, 19(2), 235–49.
- Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents, *American Economic Review*, 95(3), 616–36.
- Bolton, P., & Dewatripont, M. (2005). *Contract Theory*. Cambridge, Massachusetts: The MIT Press.
- Caloghirou, Y., Ioannides, S., & Vonortas, N. (2003). Research joint ventures. *Journal of Economic Surveys*, 17(4), 541–70.
- Calzolari, G., & Spagnolo, G. (2008). *Relational Procurement*. Retrieved 16 January 2015 from [www.csef.it/4th\\_C6/Spagnolo.pdf](http://www.csef.it/4th_C6/Spagnolo.pdf)
- Coase, R. (1937). The nature of the firm, *Economica*, 4, 386–405.
- Cremer, J., Garicano, L., & Prat, A. (2004). *Codes in organizations*. CEPR Discussion Paper No. 4205. Retrieved 27 January 2015 from <http://ssrn.com/abstract=500913>
- Dessein, W. (2002). Authority and communication in organizations, *Review of Economic Studies*, 69(4), 811–38.
- Dewatripont, M., & Tirole, J. (1999). Advocates, *Journal of Political Economy*, 107, 1–39.
- Dincer, N., & Eichengreen, B. (2014). Central bank transparency and independence: Updates and new measures, *International Journal of Central Banking*, 10(1), March, 189–253.

- Dixit, D. (2002). Incentives and organizations in the public sector, *Journal of Human Resources*, 37, 696–727.
- Domberger, S., & Rimmer, S. (1994). Competitive tendering and contracting in the public sector: A survey, *International Journal of the Economics of Business*, 1(3), 439–53.
- Duggan, M. (2000). Hospital ownership and public medical spending, *Quarterly Journal of Economics*, 115(4), November, 1343–73.
- Evans, L., & Meade, R. (2005). *The Role and Significance of Cooperatives in New Zealand Agriculture: A Comparative Institutional Analysis*. Report prepared for the New Zealand Ministry of Agriculture and Forestry, Wellington.
- Falk, M. (2006). What drives business research and development (R&D) intensity across Organisation for Economic Co-operation and Development (OECD) countries? *Applied Economics*, 38, 533–47.
- Farrell, J., & Klemperer, P. (2007). Coordination and lock-In: Competition with switching costs and network Effects, in Armstrong, M., & Porter, R. (eds.), *Handbook of Industrial Organization*, 3, 1967–2072. Amsterdam: Elsevier.
- Frey, B. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty, *Economic Inquiry*, 31, October, 663–70.
- Frey, B., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out, *American Economic Review*, 87(4), 746–55.
- Glaeser, E., & Shleifer, A. (2001). Not-for-profit entrepreneurs, *Journal of Public Economics*, 81, 99–115.
- Grossman, S., & Hart, O. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy*, 94(4), 691–719.
- Guadalupe, M., & Wulf, J. (2010). The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies, *American Economic Journal: Applied Economics*, 2(4), 105–27.
- Hart, O. (1983). The market mechanism as an incentive scheme, *Bell Journal of Economics*, 14, 366–82.
- Hart, O., & Moore, J. (1990). Property rights and the nature of the firm, *Journal of Political Economy*, 98(6), 1119–58.
- Hart, O., & Moore, J. (2005). On the design of hierarchies: Coordination versus specialization, *Journal of Political Economy*, 113(4), 675–702.
- Hart, O., Shleifer, A., & Vishny, R. (1997). The proper scope of government: Theory and an application to prisons, *Quarterly Journal of Economics*, 112(4), 1127–61.
- Heyes, A. (2005). The economics of vocation or ‘why is a badly paid nurse a good nurse’? *Journal of Health Economics*, 24, 561–69.
- Holden, R., & Posner, E. (2014). Antitrust law and internal firm efficiency. Presentation slides at the 2nd ATE symposium – Antitrust economics and competition policy: Mergers, partnerships, and innovation, University of New South Wales, Sydney. Retrieved 16 January 2015 from [http://ate.massey.ac.nz/2nd ATE Symposium files/Slides/Contributed Sessions/Richard Holden.pdf](http://ate.massey.ac.nz/2nd%20ATE%20Symposium%20files/Slides/Contributed%20Sessions/Richard%20Holden.pdf)
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics & Organization*, 7, 24–52.
- Holmstrom, B., & Tirole, J. (1991). Transfer pricing and organizational form, *Journal of Law, Economics & Organization*, 7(2), Autumn, 201–28.
- Jensen, M. (1993). The modern industrial revolution, exit, and the failure of internal control systems, *The Journal of Finance*, 48(3), July, 831–80.

- Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, 3, 305–60.
- Keller, W. (2004). International technology diffusion, *Journal of Economic Literature*, 42, 752–82.
- Kelman, S. (1990). *Procurement and public management: The fear of discretion and the quality of government performance*. Retrieved 19 August 2014 from [www.aei.org/files/2014/07/22/-procurement-and-public-management\\_170607823492.pdf](http://www.aei.org/files/2014/07/22/-procurement-and-public-management_170607823492.pdf)
- Klein, B., Crawford, R., & Alchian, A. (1978). Vertical integration, appropriable rents, and the competitive contracting process, *Journal of Law and Economics*, 21, 297–326.
- Kosfield, M., & von Siemens, F. (2011). Competition, cooperation, and corporate culture, *RAND Journal of Economics*, 42(1), Spring, 23–43.
- Kranton, R. (2003). Competition and the incentive to produce high quality, *Economica*, 70, 385–404.
- Laffont, J.-J., & Martimort, D. (2002). *The theory of incentives: The principal-agent model*. Princeton, NJ: Princeton University Press.
- Le Grand, J. (2003). *Motivation, agency, and public policy: Of knights and knaves, pawns and queens*. New York: Oxford University Press.
- Lerner, J., & Tirole, J. (2005). The economics of technology sharing: Open source and beyond, *Journal of Economic Perspectives*, 19(2), Spring, 99–120.
- Lerner, J., & Wulf, J. (2007). Innovation and incentives: Evidence from corporate R&D, *Review of Economics and Statistics*, 89, 634–44.
- Lundvall, B.-A. (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. London: Pinter.
- MacDuffie, J., & Helper, S. (2005). Collaboration in supply chains with and without trust, in Hecksher, C., & Alder, P. (eds.). *The Corporation as a collaborative community: Organization in the knowledge-based economy*. Oxford: Oxford University Press.
- Menocal, A. (2011, October). *Why electoral systems matter: An analysis of their incentives and effects on key areas of governance*. Overseas Development Institute research report. Retrieved 2 February 2015 from [www.isn.ethz.ch/Digital-Library/Publications/Detail/?lng=en&id=141037](http://www.isn.ethz.ch/Digital-Library/Publications/Detail/?lng=en&id=141037)
- Ministry of Business, Innovation and Employment. (2014). Government rules of sourcing: Rules for planning your procurement, approaching the market and contracting. Retrieved 9 February 2015 from [www.business.govt.nz/procurement/for-agencies/key-guidance-for-agencies/the-new-government-rules-of-sourcing](http://www.business.govt.nz/procurement/for-agencies/key-guidance-for-agencies/the-new-government-rules-of-sourcing)
- Moszoro, M., & Spiller, P. (2012). *Third-party opportunism and the nature of public contracts*. National Bureau of Economic Research Working Paper 18636. Retrieved 1 February 2015 from [www.nber.org/papers/w18636](http://www.nber.org/papers/w18636)
- Moszoro, M., Spiller, P., & Stolorz, S. (2014, September). The rigidity of public contracts. Retrieved 11 February 2015 from <http://ssrn.com/abstract=2469270>
- North, D. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.
- North, D. (2003). *The role of institutions in economic development*. New York and Geneva: United Nations.
- OECD. (2005). *Oslo Manual: Guidelines for collecting and interpreting innovation data*, 3rd ed. Paris: Organisation for Economic Co-operation and Development.

- Ostrom, E. (2000). Collective action and the evolution of social norms, *Journal of Economic Perspectives*, 14(3), Summer, 137–58.
- Panzar, J., & Willig, R. (1981). Economies of scope, *The American Economic Review*, 71(2) 268–72.
- Persson, T., Roland, G., & Tabellini, G. (2007). Electoral rules and government spending in parliamentary democracies, *Quarterly Journal of Political Science*, 2(2), 155–88.
- Prendergast, C. (2007). The motivation and bias of bureaucrats, *American Economic Review*, 97(1), 180–96.
- Richter, R. (2005). The New Institutional Economics: Its start, its meaning, its prospects, *European Business Organization Law Review*, 2, June, 161–200.
- Rose, N., & Joskow, P. (1990). The diffusion of new technologies: Evidence from the electric utility industry, *RAND Journal of Economics*, 21(3), Autumn, 354–73.
- Sah, R., & Stiglitz, J. (1986). The architecture of economic systems: Hierarchies and polyarchies, *American Economic Review*, 76, 716–27.
- Saussier, S., Staropoli, C., & Yvrande-Billon, A. (2009). Public-private agreements, institutions and competition: When economic theory meets facts, *Review of Industrial Organization*, 35, 1–18.
- Schmidt, K. (1997). Managerial incentives and product market competition, *Review of Economic Studies*, 64, 191–213.
- Smith, K. (2006). *Public policy framework for the New Zealand innovation system*. Ministry of Economic Development Occasional Paper 06/06. Retrieved 28 January 2015 from [www.med.govt.nz/about-us/publications/publications-by-topic/occasional-papers/2006/06-06-pdf/view](http://www.med.govt.nz/about-us/publications/publications-by-topic/occasional-papers/2006/06-06-pdf/view)
- Spiller, P. (2008, August). *An institutional theory of public contracts: Regulatory implications*. NBER Working paper 14152. Cambridge: Massachusetts: National Bureau of Economic Research.
- Spiller, P., Stein, E., & Tommasi, M. (2003). *Political institutions, policymaking processes, and policy outcomes: An intertemporal transactions framework*. Retrieved 30 January 2015 from [www.yale.edu/leitner/resources/docs/2003-03.pdf](http://www.yale.edu/leitner/resources/docs/2003-03.pdf)
- Tirole, J. (1988). *The theory of industrial organization*. Cambridge, Massachusetts: MIT Press.
- Williamson, O. (1999). Public and private bureaucracies: A transaction cost economics approach, *Journal of Law, Economics & Organization*. 15(1), 306–42.
- Williamson, O. (2000). The New Institutional Economics: Taking stock, looking ahead, *Journal of Economic Literature*, 38, 595–613.
- Williamson, O. (2002). The theory of the firm as governance structure: From choice to contract, *Journal of Economic Perspectives*, 16(3), Summer, 171–95.
- Winter, E. (2004). Incentives and discrimination, *American Economic Review*, 94(3), 764–73.
- Yu, D., & Hang, C. C. (2009). A reflective view of disruptive innovation theory, *International Journal of Management Reviews*, 1–21.