WPS-2395

Decentralizing the Provision of Health Services

How should central and local governments allocate authority for the planning, financing, and delivery of health services?

An Incomplete Contracts Approach

William Jack

Summary findings

Jack studies the allocation—between a central government and a local authority—of responsibility for planning, financing, and operations for the delivery of health services, in the context of an incomplete contracts model. In this model, inputs are required of both the central government and local authorities but they are unable to write down, and commit to, a complete and binding contract describing the actions both should take.

The model is meant to capture the tradeoff between central and local authority in decisions about both financing and the provision of services. Each party provides a specific input—for example, the central government establishes a drug procurement system while the local authority designs and implements an incentive scheme to get doctors to carry out their responsibilities appropriately. The responsibility for delivery of services is identified with the ownership of essential infrastructure, such as the clinic or hospital.

Jack finds that to maximize the joint surplus of the two public bodies:

• Ownership of the facility should be given to the party that most values the well-being of local residents.

(This way, if ex post bargaining breaks down, each still enjoys some benefits from the other's actions.)

• Financing authority and responsibility for delivering services should be negatively correlated. Generally it is optimal to allocate tax authority to the party that values the residents' well-being less—in other words, separate spending responsibility (ownership) from financing authority. A heavier financing burden (access to a small and inefficient tax base) has the same incentive effect as asset ownership: It increases the return to effort.

• If transferring ownership of the physical asset is costly (because the party that builds the asset has an inherent advantage in operating it—that is, there is some human capital embodiment), it may be optimal for the party with the higher construction costs to have planning authority.

• Somewhat paradoxically, the greater the costs of transferring assets from one party to the other, the more likely that ownership of the facilities and their provision should be separated.

Decentralizing the Provision of Health Services

An Incomplete Contracts Approach[1]

William Jack

Consultant, Development Research Group, World Bank

1818 H Street NW, Washington, DC 20433, USA

1 Introduction

Poor countries face obvious resource constraints on all dimensions of consumption and investment, including those in the health sector. However, by the same token, estimates have suggested that relatively small resource allocations could have not inconsiderable effects on health outcomes and welfare. For example, the World Bank (1993, p. 106) estimated a basic package of care would cost low income developing countries in the vicinity of US$4 per capita (1.2% of GDP), and would cost middle income countries US$7 per capita (just 0.3% of GDP).

If effective health care is so cheap, why has so little progress been made in addressing the appalling living conditions and health status of poor people in the developing world? The answer must be that the resources available are not used efficiently or equitably, despite (or sometimes because of) the objectives of government. Of course, it has long been well established that the organization of the delivery of health care is highly non-trivial, as the often delicate balance between equity and efficiency concerns under conditions of risk, market failure, and asymmetric information is struck. Indeed, the question of why public spending has been relatively ineffectual in improving health indicators in developing countries has recently been empirically investigated by Filmer, Hammer, and Pritchett (1998), who identify incentive constraints on the public sector's ability to provide services.[1]

One response that has been adopted recently in many developing countries, often with World Bank support, is to decentralize the provision of health (and other social) services, in the hope that local governments will use resources more efficiently. In particular, it is hoped that lower level autonomy over health care provision can lead to enhanced allocative efficiency (better matching services provided to those required/demanded) and to greater production efficiency, such as by ensuring compliance by providers with government requirements (e.g., showing up for work). This paper examines a simple application of recent ideas from the theory of the firm to the question of

---

[1] They also identify the possibility that public spending might crowd out private supply, so the net effect on utilization can be small.

the allocation of authority in the delivery of health services.

The model I will write down here is meant to capture the trade-off between central and local authority over decisions regarding both provision of services and financing. I assume that both parties - the central and local governments - have something to contribute to the efficient provision of health services. Thus even in a decentralized regime, the centre has some role to play (organizing national coordination, providing national public good inputs into health production, acting as a potential source of financing); and conversely, in a centralized arrangement, the local authority is active, perhaps acting as a monitor and/or regulator of physician behavior.

Orthodox discussion of decentralization concentrates on things like interjurisdictional externalities and national public goods, versus local information about preferences and production techniques (Oates (1999)). It can be argued that these characteristics do not get to the heart of the decentralization question, which is about how decision authority should be allocated. Externalities can be fixed with taxes and subsidies, or through Coasean bargaining, for example, and information asymmetries point to delegation of tasks, but not necessarily decentralization of authority. Some of the more formal analysis of decentralization has employed the principal-agent paradigm in analyzing the design of contractual relationships between central and local government. There are a number of shortcomings with this approach however. First, the allocation of tasks between the two parties is exogenous,[2] so a principal question of decentralization - who should do what - is difficult to address in this framework. Second, by their nature, many of the tasks that are to be performed, and the underlying objectives, are difficult to describe, particularly in the social sectors. This makes writing formal contracts between central and local governments difficult at best. And third, it might be relatively easy for a central government to renege on specific aspects of a contract (e.g., revenue sharing formulae), yielding finely-tuned contracts open to future renegotiation. I will assume that more robust aspects of intergovernmental relations (in particular asset ownership and tax base authority)

---

[2] Holmstrom and Milgrom (1987 and 1991) examine which tasks an agent should be required to do (or prohibited from doing), but do not assign tasks between principal and agent.

are less susceptible to arbitrary alteration by the central powers.[3]

Within the realm of principal-agent models, probably the most useful question to ask is not what the optimal contract between the principal (central government) and agent (local authority) should be, but more basically who should the principal be? That is, is it better to have a central government bureaucrat, answerable to a central ministry, monitor the behavior of a service provider, or to have a local politician answerable to the local population perform this function? This is the issue of accountability. The answer will depend on the importance of externalities etc., and on the extent to which the local population can control its politicians (strength of local democracy, likelihood of capture, etc.). Accountability to the population is somewhat more indirect in the centralized case (the bureaucrat is accountable to the minister, who, as part of the government, is accountable to the national population) than in the decentralized case, although it is not clear a priori that there is more accountability under the latter regime.[4]

One can understand issues of accountability as determining the extent to which the objectives of local and central governments coincide with some notion of aggregate social welfare, as judged by an objective outside observer. Whatever the degree of convergence of these objectives, there remains the task of coordinating the actions of the two parties. This is the subject of this paper.

Given these observations and concerns, I will employ a simple incomplete contracts model of decentralization. I borrow extensively from Hart's (1995) model of vertical integration, although there are some differences that arise due to the fact that health outputs represent a public good that is valued by both levels of government. The essence of Hart's analysis is that, absent formal incentive contracts, ownership of assets used in production can provide incentives, and the allocation of ownership rights determines aggregate surplus, through its effects on incentives. In turn, these

---

[3] This requires some kind of well-functioning court system. Governments can and do nationalize whole industries, often counter to guarantees of non-intervention. In the health sector this would appear to be less likely, as it is usually not a great revenue generator.

[4] See Seabright (1998) for an incomplete contracts model of accountability within a decentralized economy. Also Bardhan and Mookherjee (1998) use an incomplete contracts approach to model local capture of politicians.

incentives derive from the inevitable ex post renegotiation of any existing contract between two parties, and the fact that outside options (which are a function of asset holdings) determine each party's bargaining power. In the context of the organization of government provided services, the allocation of decision making authority (which is like asset ownership) determines incentives and the net equilibrium welfare improvement, or quality of care.

More specifically, I will assume that production of medical care and related services requires three invested (i.e., sunk) inputs. The first can be thought of as a physical capital good, $K \geq 0$, such as a clinic/hospital and associated equipment. This capital good can be provided by either the central government (G) or the local authority (A). The other two inputs are $e_G \geq 0$ and $e_A \geq 0$, and can be provided only by the central government and the local authority, respectively. e_G is something like a drug procurement and distribution network, and e_A is a local organization and monitoring system for inducing physicians to perform.[5] A time line is shown in Figure 1.

There is no question, in this model, about which party should perform the functions e_G and e_A. The questions I want to address are (i) who should own the facility, K, (ii) who should have financing authority, and (iii) who should build the facility? There are thus three types of decentralization possible in this model, depending on the answers to these questions: decentralized ownership, fiscal decentralization (specific to the project), and decentralized provision. One of the contributions of the paper is to examine the conditions under which these three aspects of decentralization should go together. For example, ownership of K can provide incentives for provision of e_G and e_A. Because of this, the identities of the owner and builder of K may differ, and it is not always the case that the most efficient provider of the physical asset should build it. This helps formalize ideas of participation and the importance of "local ownership" in project design, and helps identify the trade-offs between such benefits and the associated costs.

[5] In the standard theory of the firm model, e_G and e_A would be interpreted as human capital investements by each of two interacting parties. The essential feature of e_G and e_A is that their ownership cannot be allocated arbitrarily to one party or the other, and that only the party that invests in each can "own" it.

```
┌─────────────────────────────────────┐
│ Allocate responsibilities and rights │
│ (i) Hospital built by i              │
│ (ii) Hospital owned by j             │
│ (iii) Tax authority given to k       │
└─────────────────────────────────────┘
                  ↓
        ┌──────────────────┐
        │ i builds hospital │
        └──────────────────┘
                  ↓
        ┌──────────────────────────┐
        │ Ownership transferred to j │
        └──────────────────────────┘
                  ↓
┌──────────────────────────────────────────┐
│ G develops drug procurement system        │
│ A develops local monitoring and incentive system │
└──────────────────────────────────────────┘
                  ↓
┌──────────────────────────────────────────┐
│ G and A negotiate on how to combine inputs │
│ and how to finance provision              │
└──────────────────────────────────────────┘
                  ↓
┌──────────────────────────────────────────┐
│ Joint surplus realized and divided         │
│ according to ex ante bargaining power      │
└──────────────────────────────────────────┘
```
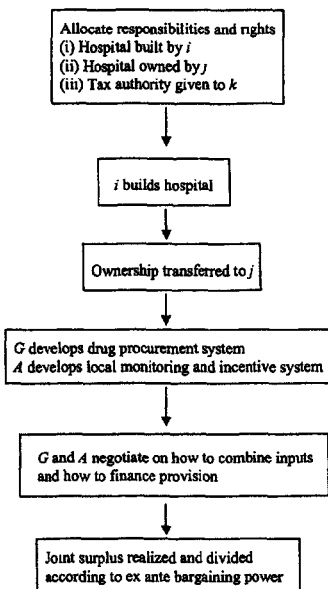
Figure 1:

The analysis can be done in two parts. First, I assume the existence of a physical asset, K, and ask who should own it, and how financing should be arranged. Then, working backwards, I address the question of who should build it. The next section sets up the structure of the model. Section 3 looks at ownership of K, the following section looks at tax assignment within the incomplete contract model. Section 5 simplifies the underlying structure of the model in order to examine who should provide K, and section 6 draws some conclusions.

2 Technology and Preferences

2.1 Quality

The three assets K, e_G, and e_A can be combined with fixed supplies of other inputs (e.g., doctors, drugs, etc.) to produce health services.[6] If they are used jointly and cooperatively by the two parties, then the quality of services is $\Psi(e_G, e_A, K)$. On the other hand, if the parties do not cooperate, then

---

[6] Having fixed supplies of recurrent inputs makes it easy to talk about quality and cost differences independently.

5

ownership of K matters. For example, if G owns the physical asset, then it can try to provide medical care using K and e_G, but not e_A. At the same time, A can provide some services using just the input e_A. We denote the net effect of these two independent means of provision by $\psi^G(e_G, e_A, K)$. If on the other hand the local authority owns the physical asset, then under non-cooperation it can use K and e_A to produce services, while G uses just e_G, yielding net quality $\psi^A(e_G, e_A, K)$. Ψ, ψ^G, and ψ^A are assumed to be increasing and concave in each of their arguments.

We also assume that access to the asset K improves the absolute and marginal productivity of efforts. That is, for $i = G$ and A,

$$\Psi(e_G, e_A, K) > \psi^i(e_G, e_A, K),$$

and

$$\Psi_1(e_G, e_A, K) > \psi_1^G(e_G, e_A, K) > \psi_1^A(e_G, e_A, K), \text{ and}$$

$$\Psi_2(e_G, e_A, K) > \psi_2^A(e_G, e_A, K) > \psi_2^G(e_G, e_A, K).$$

where the subscript denotes the partial derivative with respect to the relevant argument. To ensure existence of interior optima, I assume also that $|\Psi_{jj}| > |\psi_{jj}^i|$ for $j = 1, 2$.

2.2 Cost

Using the assets K, e_G and e_A is costly (drugs have to be purchased, doctors paid, and buildings maintained). Cooperation has an effect on these costs of provision just as it does on quality. Under cooperation, operating costs are given by the function $\Gamma(e_G, e_A, K)$. If the parties do not cooperate, each operates those assets under its control separately, and total costs are

$$\gamma^G(e_G, e_A, K) = \eta_G(e_G, K) + \eta_A(e_A, 0)$$

if G owns the facility, and

$$\gamma^A(e_G, e_A, K) = \eta_A(e_A, K) + \eta_G(e_G, 0)$$

6

if A does. η_i is the stand-alone cost that party i must finance in operating the assets it has under its control. It is assumed that effort and access to the physical asset both reduce a party's stand-alone costs, i.e., that[7]

$$\frac{\partial \eta_i}{\partial e_i} < 0 \text{ and } \frac{\partial \eta_i}{\partial K} < 0.$$

In addition, it is assumed that the productivity of effort in reducing costs is increasing in the level of the physical asset available, i.e.,

$$\frac{\partial^2 \eta_i}{\partial e_i \partial K} < 0.$$

In particular, $|\partial \eta_i(e_i, K)/\partial e_i| > |\partial \eta_i(e_i, 0)/\partial e_i|$, so owning the asset makes own effort more productive.

We again assume that cooperation is good in both senses:

$$\Gamma(e_G, e_A, K) < \gamma^i(e_G, e_A, K)$$

and

$$|\Gamma_j(e_G, e_A, K)| > |\gamma^i_j(e_G, e_A, K)|.$$

Again, to ensure existence, it is assumed that $|\Gamma_{jj}| > |\gamma^i_{jj}|$. Thus cooperation lowers operating costs, and increases the cost reduction associated with an increase in each of the assets K, e_G, and e_A. For example, a higher level of e_G might correspond to a better drug procurement and management system that reduces drug waste. This would mean that fewer drugs would need to be purchased to attain a given quality of service (or that their prices would fall), lowering operating costs.

2.3 Consumer welfare

Quality is taken as a proxy measure for the benefits consumers receive from the services provided. I assume also that the consumers finance the costs of provision through tax payments, (possibly

---

[7] Recall that these costs represent the costs of other inputs purchased in order to provide a fixed level of services. It is therefore uncontroversial that higher levels of inputs of effort and K reduce the requirements for other inputs, and hence the unit cost that needs to be financed.

including user fees). The two levels of government have different taxes bases, the marginal distortionary costs of which are λ_G and λ_A respectively. Under cooperation, the more efficient tax base is used. Consumer well being in this case is

$$W^*(e_G, e_A, K) = \Psi(e_G, e_A, K) - (1 + \lambda_{\min})\Gamma(e_G, e_A, K) \tag{1}$$

where $\lambda_{\min} = \min\{\lambda_G, \lambda_A\}$. In the non-cooperative regimes, consumer well-being is

$$W^G(e_G, e_A, K) = \psi^G(e_G, e_A, K) - [(1 + \lambda_G)\eta_G(e_G, K) + (1 + \lambda_A)\eta_A(e_A, 0)] \tag{2}$$

under G-ownership, and

$$W^A(e_G, e_A, K) = \psi^A(e_G, e_A, K) - [(1 + \lambda_A)\eta_A(e_A, K) + (1 + \lambda_G)\eta_G(e_G, 0)] \tag{3}$$

under A-ownership.

Our assumptions above on the effects of K on the productivity of effort in reducing stand-alone costs imply that

$$W_1^*(e_G, e_A, K) > W_1^G(e_G, e_A, K) > W_1^A(e_G, e_A, K) \tag{4}$$

and

$$W_2^*(e_G, e_A, K) > W_2^A(e_G, e_A, K) > W_2^G(e_G, e_A, K). \tag{5}$$

2.4 Government objectives

The two governments care about two things: (i) the well-being of consumers (although to possibly different extents), and (ii) the non-pecuniary costs associated with their investment choices (effort costs).[8] Thus, G's objective, given K is

$$u_G(e_G, e_A, K) = \beta_G W(e_G, e_A, K) - e_G$$

---

[8] I abstract from issues of inter-jurisdictional externalities, that may be internalized by the central government but not by the local authority.

where $W(.) = W^*(.)$, $W^G(.)$, or $W^A(.)$. Similarly, A's objective is

$$u_A(e_G, e_A, K) = \beta_A W(e_G, e_A, K) - e_A.$$

We normalize the parties' preference weights so that $\beta_G + \beta_A = 1$, and interpret $\beta_G > \beta_A$ as describing a situation in which the local authority is less responsive to residents' desires than is the central government, say due to capture by the local elite. On the other hand, $\beta_G < \beta_A$ represents the case where a far away central government is less responsive to local needs.

Having supplied efforts e_G and e_A, the total surplus generated is $\beta_G W + \beta_A W$. Under cooperation, this is just $W^*(e_G, e_A, K)$, while under non-cooperative i-ownership it amounts to $W^i(e_G, e_A, K)$, for $i = G$ and A.

3   Optimal asset ownership

Given the existence of a physical asset, K, who should own it? To answer this we just compare the gross surplus (as calculated by the two governments) that will be generated under the two possible ownership structures - G-ownership and A-ownership.

If the parties could contract on e_G and e_A, then the first-best outcome could be achieved - that is, e_G and e_A would be chosen to maximize W^*, and this aggregate surplus would be shared between the two in some fashion depending on the division of ex ante bargaining power. However each party recognizes that when such contracts are not feasible (or enforceable), then ex post bargaining will mean that at least some of the returns to its effort will be shared with the other player. This reduces incentives to provide effort, and the optimal ownership allocation is such as to minimize these distortions. We take the standard approach and assume that ex post Nash bargaining results in a 50:50 split of the gains from cooperation.

Exactly how the implied transfer from one party to the other associated with splitting the gains from trade is effected is not modeled here explicitly. One possibility is that this transfer is implemented implicitly through adjustments to existing inter-governmental grants, or that responsibility

9

for provision of other public services is reallocated between the parties.

3.1 G-ownership

3.1.1 Payoffs

Under G-ownership, suppressing all functional dependencies for convenience, the gains from trade are

$$X^G = W^* - W^G$$

At the ex post bargaining stage, effort costs are sunk, so the payoff to G is thus

$$\begin{aligned} \pi_G^G &= \beta_G W^G + \frac{1}{2}X^G \\ &= \frac{1}{2}W^* + \frac{1}{2}\Delta\beta W^G \end{aligned}$$

where $\Delta\beta = \beta_G - \beta_A$, and we have used $\beta_G + \beta_A = 1$. The first term in the first line is G's payoff if bargaining breaks down.[9] Similarly, A's ex post payoff under G-ownership is

$$\begin{aligned} \pi_A^G &= \beta_A W^G + \frac{1}{2}X^G \\ &= \frac{1}{2}W^* - \frac{1}{2}\Delta\beta W^G \end{aligned}$$

Notice that if both governments give consumer welfare the same weight ($\beta_G = \beta_A$), then $\pi_G^G = \pi_A^G = \frac{1}{2}W^*$, reflecting that any benefit that accrues to consumers represents a pure public good vis-à-vis the two governments.

3.1.2 Effort choices

The ex interim payoff to G under G-ownership is $\pi_G^G - e_G$,[10] and its choice of effort satisfies

$$\frac{1}{2}W_1^* + \frac{1}{2}\Delta\beta W_1^G = 1. \tag{6}$$

---

[9] Notice that we have assumed that even if bargaining breaks down, so that both parties provide services independently, Gs utility depends on the total quality delivered to, and costs imposed on, residents, ψ^G and γ^G, respectively, and not just on the quality it provides and the costs it imposes.

[10] Ex interim because K has been sunk at this stage.

10

Similarly, A's payoff under G-ownership is $\pi_A^G - e_A$, and its effort choice satisfies

$$\frac{1}{2}W_2^* - \frac{1}{2}\Delta\beta W_2^G = 1. \tag{7}$$

Note that since $|\Delta\beta| < 1$, our convexity assumptions on the Ψ/ψ and Γ/γ functions guarantee that the second order conditions are both satisfied.[11] Denote the optimal effort choices of G and A by $\hat{e}_G^G$ and $\hat{e}_A^G$ respectively. These effort choices will in general be functions of the size (quality) of the existing physical capital asset, K, but we suppress such dependence for now. When $\beta_G = \beta_A$, (6) and (7) simplify to

$$\frac{1}{2}W_1^* = 1 \tag{8}$$

and

$$\frac{1}{2}W_2^* = 1. \tag{9}$$

In this case, although the resulting effort choices are not first-best (that would require $W_i^* = 1$, for $i = 1, 2$), they are independent of asset ownership.

The total ex post surplus thus generated from G-ownership is $W^*(\hat{e}_G^G, \hat{e}_A^G, K)$ - that is, effort levels will be chosen sub-optimally, but given these efforts, the parties will cooperate (so value the assets using $W^*(.)$). Let us denote this surplus, as a function of K only, by $S^G(K) = W^*(\hat{e}_G^G, \hat{e}_A^G, K)$.

3.2 A-ownership

Under A-ownership, payoffs are simply

$$\pi_G^A = \frac{1}{2}W^* + \frac{1}{2}\Delta\beta W^A$$

for G and

$$\pi_A^A = \frac{1}{2}W^* - \frac{1}{2}\Delta\beta W^A$$

---

[11] The second order conditions are $W_{11}^* + \Delta\beta W_{11}^G < 0$ and $W_{22}^* - \Delta\beta W_{22}^G < 0$ respectively.

11

for A, yielding optimal effort choices $\widehat{e}_G^A$ and $\widehat{e}_A^A$ that satisfy

$$\frac{1}{2}W_1^* + \frac{1}{2}\Delta\beta W_1^A = 1 \qquad (10)$$

and

$$\frac{1}{2}W_2^* - \frac{1}{2}\Delta\beta W_2^A = 1. \qquad (11)$$

The total surplus thus generated from A-ownership is $S^A(K) = W^*(\widehat{e}_G^A, \widehat{e}_A^A, K)$.

3.3 Comparison with private goods model

Let us consider the first order condition for G's effort choice under G-ownership, (6), and write it as

$$\frac{1}{2}W_1^* + \frac{1}{2}\beta_G W_1^G - \frac{1}{2}\beta_A W_1^G = 1.$$

In Hart's (1995) basic model, only the (equivalents of the) first two terms appear on the left hand side. This is because if bargaining breaks down in that model, the effort of one party has no effect on the welfare (profit) of the other. However here, even when bargaining breaks down, the public good nature of each party's action means that the other's utility is affected. For example, G makes its choice assuming it gets half the increase in the gains from trade - but these gains are reduced not only by the increase in its own outside option, but also through the increase in A's outside option when e_G increases. This helps explain also the fact that when $\beta_G = \beta_A$ asset ownership has no effect on incentives to provide effort. In this case, the increase in G's outside option resulting from an increase in its effort is exactly offset by a reduction in the gains from trade due to the concomitant increase in A's outside option.

3.4 Optimal ownership allocation

For a fixed level of K, the optimal ownership structure is that which maximizes the total surplus. That is, G-ownership is optimal if and only if $S^G(K) > S^A(K)$. Consider then the case where the central government places higher weight on the well-being of residents than does the local authority,
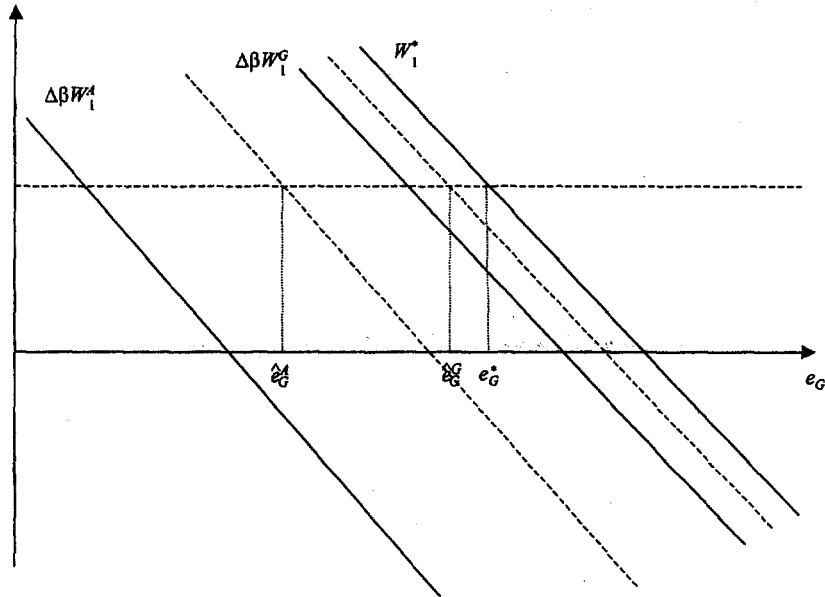
Figure 2:

$\Delta\beta > 0$. Figure 2 shows the optimal effort choice by G under the two ownership regimes: in each case effort is below the first best level, e_G^*, but the marginal productivity of e_G is higher under G-ownership than under A-ownership, in accordance with condition (4).

Figure 3 similarly shows A's effort choice under the different regimes, again under the assumption that the capital asset increases the productivity of e_A (condition (5)). As drawn, it is clear that despite the positive effect of ownership on the productivity of effort, A should definitely not be given ownership rights, as this reduces its effort choice to $\hat{e}_A^A$, which is well below $\hat{e}_A^G$. Indeed, as long as $\Delta\beta > 0$, A's effort choice will always be higher under G-ownership than under A-ownership. To ensure that A's effort in such cases is below the first-best level, the following assumption is made: Let $\tilde{e}_A^X$ satisfy $W_2^*(\hat{e}_G, e_A, K) = -W_2^X(\hat{e}_G, e_A, K)$; then $W_2^*(\hat{e}_G, \tilde{e}_A^X, K) < 1$.

We can now make the following observations:

Proposition 1 *If $\Delta\beta > 0$, G's effort is always larger when it owns the asset, and A's is always smaller when it does.*
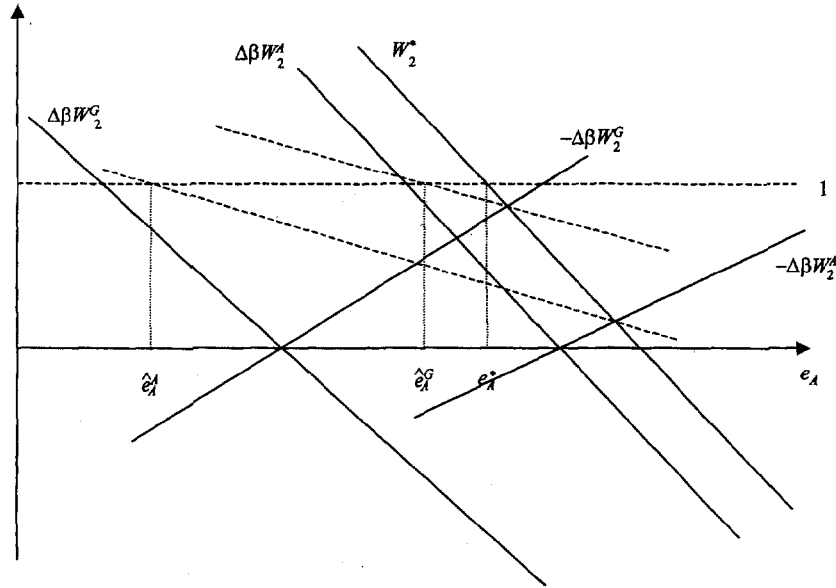
13

Figure 3:

Proposition 2 *Conversely, if $\Delta\beta < 0$, G's effort is always smaller when it owns the asset, and A's is always larger when it does.*

Proposition 3 *Since the two effort levels are always below their first best levels, it is optimal to assign ownership of the physical asset to the party that puts a higher weight on the well-being of local residents.*

A number of points regarding these results are worth noting. First, recall that ownership allocation is used to maximize the joint surplus of the two contracting parties, that is, the local authority and the central government, and the well-being of local residents only matters (in this model) to the extent that it enters the objective functions of these parties. Thus the final result above should not necessarily be interpreted as deriving from an external observer's desire to maximize residents' well-being subject to the actions of the two parties.

Second, suppose G values residents' well being more, so $\Delta\beta > 0$. Then even when G's effort is relatively useless unless it is combined with both e_A and K, it is still the case that it should own

14

the asset. Giving the asset to G stops A from choosing the very low effort level $\widehat{e}_A^A$.

Finally, as a generalization of the forgoing remark, the allocation of ownership in this model is less sensitive to technological parameters than in the standard private goods model of Hart (1995). When there are no external effects of effort choice in the event that bargaining breaks down, asset ownership in general is associated with the impact of such on the productivity of effort. Here however, no matter how small the impact of ownership on the marginal productivity of effort, asset ownership is a function only of the underlying preference parameters. This suggests a certain robustness to the ownership assignment rule.

4   Allocation of tax authority and financing

Decentralization is often a thinly veiled attempt by the central government to improve its budgetary position, by transferring obligations for provision of services without matching these with sufficient funds. There is generally a presumption that responsibility for provision should be matched by adequate finance and/or responsibility to tax. The model in this paper can shed light on the role of tax base assignment - or more generally, the shadow cost of public funds - on the allocation of responsibility/ownership.

Let us consider then the effect of alternative tax powers on equilibrium effort provision, taking as given the allocation of ownership of the capital asset, K. Assume for concreteness that $\Delta\beta > 0$, so the central government places a higher value on the well-being of local residents. Figures 4 and 5 show the equilibrium effort choices, $\widehat{e}_G^X$ and $\widehat{e}_A^X$, where X is the (for now fixed) owner of the capital asset. The positions of the curves W_1^X and W_2^X are determined by, inter alia, the values of λ_G and λ_A, as seen by inspection of equations (2) and (3). Because additional effort reduces financeable costs, the marginal return to effort increases as λ increases.

Starting from the position in Figure 4, consider a transfer of tax authority from the central government to the local authority that increases λ_G and reduces λ_A. Assuming that both efforts $\widehat{e}_G^X$
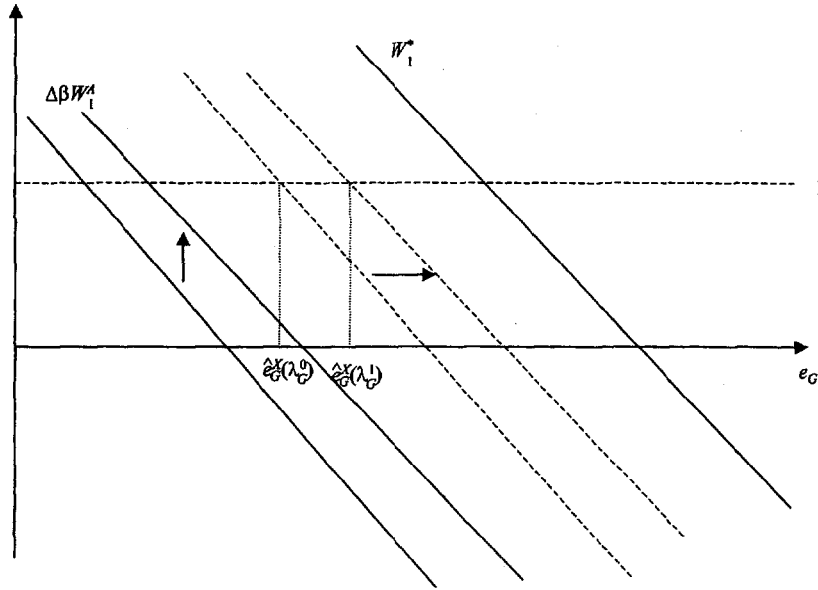
Figure 4:

and $\hat{e}_A^X$ were initially below their first-best levels, this transfer of authority unambiguously increases the effort of each party. Intuition for this result can be seen from returning to equations (6) and (7) and noting that each party's effort increases (i) as the extent to which its valuation of local residents' welfare exceeds the other's increases, and (ii) as the marginal effect of its effort increases $(W_j^X,\ j = 1, 2)$. When $\Delta\beta > 0$, for G an increase in the marginal effect of effort is weighted positively, and so its effort increases. For A, the transfer of tax authority reduces the marginal effect of effort on non-cooperative welfare (W^X), but since this is weighted negatively, the equilibrium response of A is to also increase effort. Clearly, if $\Delta\beta$ were less than zero, the same change in tax assignment would reduce both parties' equilibrium efforts. Thus we have the following result:

Proposition 4 *Tax authority should be given to the party that values the well-being of individuals relatively less.*

Coupled with our earlier result, this proposition asserts that ownership and tax authority should *not* be allocated to the same party. Such a prescription is counterintuitive because we might think
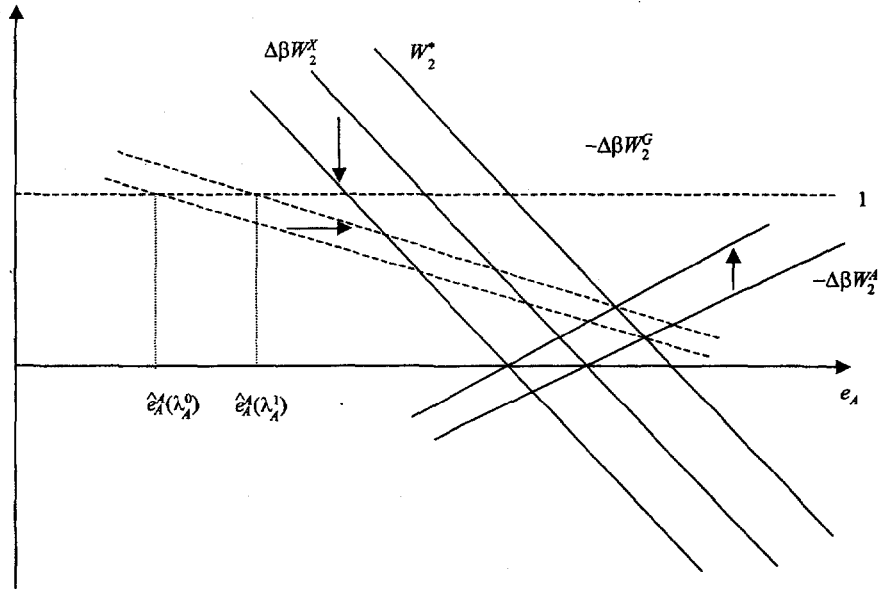
16

Figure 5:

that resources for provision should be given to the party that cares more about quality. But of course, allocation of tax authority only affects ex ante effort choices, since all financing is agreed upon cooperatively at the ex post stage of the game.

Identifying ownership of the physical asset with "expenditure responsibility" again highlights the surprising nature of this result. It is common for commentators to call for decentralization of expenditure responsibility to be matched with the allocation of a revenue source, but the analysis above suggests the opposite. Again, this can be understood by recalling that we have assumed that in equilibrium the parties successfully negotiate and use the most efficient tax source. The allocation of tax authority affects incentives to provide effort only to the extent that it influences the disagreement payoffs to each party, which are not realized in equilibrium.

5   Provision and ownership of the health care facility

We are now in a position to step back and address questions about the provision of K as well as its ownership. This question is at the heart of Grossman and Hart's (1986) earlier work, and is not unrelated to more recent analyses of privatization (Hart, Shleifer, and Vishny (1998)) and the role of NGOs (Besley and Ghatak (1999)). I will make the assumption that at the ex interim stage, after the asset has been created, the parties cannot restrain themselves from reallocating ownership of the asset if this is Pareto efficient. That is, I will assume that ex interim considerations determine the equilibrium ownership allocation, and this cannot be altered by a binding assignment at the ex ant stage. The essential question then is, given the optimal identity of the owner of the physical asset, who should produce it?

At first sight, it would appear that the party with the lower costs of production should build the facility, and indeed this is true if the parties can write a binding contract specifying the level of investment. However, if the investment level is not contractible, then assigning responsibility for provision is the only lever the parties have for inducing provision.

An important possibility that has been introduced in related literature is that the productivity of the physical asset may fall when transferred from one party to the other. For example, some authors (e.g., Hart, Shleifer, and Vishny (1998) and Besley and Ghatak (1999)) have allowed the possibility that part of the value of the physical asset be embodied in the human capital of the provider. In such cases, if the productivity of the asset is reduced enough when it changes hands, then it would appear that assigning provision responsibility to the prospective owner is appropriate. However, this assumes that the level of investment is the same under both assignments of responsibility. It will be shown in fact that as the cost of transferring the asset increases, it becomes more likely that provision by the higher cost non-owner yields a higher investment level. If the benefit of a an expanded facility outweighs the additional cost, this may be identified as a potential benefit of "participation" at the project development (i.e., ex ante) stage by an otherwise inefficient party.

To model these issues, let us assume that it is costly to have different parties produce and own the capital asset, because the asset's productivity falls when it is not used by, or with the cooperation of, the provider. Thus, suppose that if the owner is not the provider of the asset, then in the event that bargaining breaks down and the parties behave non-cooperatively, the net surplus generated is reduced by a fraction $(1 - \delta) < 1$.[12]

For the purposes of exposition, it is useful to impose a little more structure on the welfare measures in equations (1) through (3). Specifically, let us assume that benefits and costs are separable in K and the two effort inputs, and that different ownership structures only affect the impact of effort. That is, we assume (1) can be written

$$W^*(e_G, e_A, K) = \zeta_G^* e_G + \zeta_A^* e_A + \phi(K)$$

where $\phi(.)$ is increasing and concave. Similarly, (2) and (3) are assumed to reduce to

$$W^G(e_G, e_A, K) = \zeta_G^G e_G + \zeta_A^G e_A + \alpha^G \phi(K)$$

and

$$W^A(e_G, e_A, K) = \zeta_G^A e_G + \zeta_A^A e_A + \alpha^A \phi(K)$$

where $\zeta_G^* > \zeta_G^G > \zeta_G^A$, $\zeta_A^* > \zeta_A^A > \zeta_A^G$, $\alpha^G < 1$ and $\alpha^A < 1$. The idea is that just having the asset available affects the productivity of efforts, but the size of the asset does not. This is an extreme assumption, but it means that I can ignore the effect of asset size on the ex interim effort of each party, which simplifies the algebra. I also ignore financing issues, by assuming a lump-sum revenue source. Four cases can then be analyzed in turn.

G provision and G ownership

---

[12] A perhaps more appealing assumption might be that the effect of human capital embodiment is that when the asset is transferred from provider to owner, it is as if the owner is working with an asset whose size is reduced by fraction $(1 - \lambda) < 1$. That is, when the provider transfers an asset of size K, the new owner has λK units to work with. However, the algebra is a little simpler with the assumption that the effect of the transfer is on the surplus generated, and not just on the input, K.

19

G's ex ante payoff is given by

$$\rho_G^G(K) = \frac{1}{2}W^*(e_G, e_A, K) + \frac{1}{2}\Delta\beta W^G(e_G, e_A, K) - e_G - \theta_G K,$$

yielding choice of investment K_G^G satisfying[13]

$$\frac{1}{2}W_3^* + \frac{1}{2}\Delta\beta W_3^G = \theta_G,$$

which reduces to

$$\frac{\phi'(K)}{2} = \frac{\theta_G}{(1 + \Delta\beta\alpha^G)} \equiv z_G^G.$$

G provision and A ownership

G's ex ante payoff is given by

$$\rho_G^A(K) = \frac{1}{2}W^*(e_G^0, e_A^0, K) + \frac{1}{2}\Delta\beta\delta W^A(e_G^0, e_A^0, K) - e_G^0 - \theta_G K,$$

yielding choice of investment K_G^A satisfying

$$\frac{1}{2}W_3^* + \frac{1}{2}\Delta\beta\delta W_3^A = \theta_G$$

or

$$\frac{\phi'(K)}{2} = \frac{\theta_G}{(1 + \Delta\beta\delta\alpha^A)} \equiv z_G^A. \tag{12}$$

A provision and G ownership

A's ex ante payoff is given by

$$\rho_A^G(K) = \frac{1}{2}W^*(e_G^0, e_A^0, K) - \frac{1}{2}\Delta\beta\delta W^G(e_G^0, e_A^0, K) - e_A^0 - \theta_A K,$$

yielding choice of investment K_A^G satisfying

$$\frac{1}{2}W_3^* - \frac{1}{2}\Delta\beta\delta W_3^G = \theta_A$$

---

[13]Because of the additive separability, there are no $\partial\widehat{e}_X/\partial K$ terms.

or

$$\frac{\phi'(K)}{2} = \frac{\theta_A}{(1 - \Delta\beta\delta\alpha^G)} \equiv z_A^G.$$ (13)

A provision and A ownership

A's ex ante payoff is given by

$$\rho_A^A(K) = \frac{1}{2}W^*(e_G^0, e_A^0, K) - \frac{1}{2}\Delta\beta W^A(e_G^0, e_A^0, K) - e_A^0 - \theta_A K,$$

yielding choice of investment K_A^A satisfying

$$\frac{1}{2}W_3^* - \frac{1}{2}\Delta\beta W_3^A = \theta_A.$$

This simplifies to

$$\frac{\phi'(K)}{2} = \frac{\theta_A}{(1 - \Delta\beta\alpha^A)} \equiv z_A^A.$$ (14)

As discussed above, I will make the assumption that it is not possible for the parties to commit to an ownership allocation at the ex ante stage of the relationship that they will mutually wish to alter at the ex interim stage. That is, I take the allocation of ownership of the asset as being determined by the need to provide incentives for the parties to engage in effort at the ex interim stage, as determined by the sign of $\Delta\beta$. It is then straightforward to identify the optimal provider of the capital good, conditional on this ownership structure.

For clarity, let us assume without loss of generality that $\Delta\beta > 0$. This does not necessarily mean that the central government should own the facility at the ex interim stage, since this will depend on the identity of the provider. In particular, if G provides the facility, then effort incentives are strongest when it retains ownership. However, when A provides the facility, it is optimal for ownership to be transferred to G if and only if the costs of transferal are small enough. Specifically, the asset should be owned by G, despite being provided by A, if and only if

$$\delta > \frac{\zeta_G^A}{\zeta_G^G} \equiv \zeta^*.$$

21

When $\delta < \zeta^*$, the cost of transfering the facility is high enough that the improved incentives that would otherwise exist do not materialize.

With this determination of ex interim ownership we can consider the effects of different planning provision. If asset transfer is costly, $\delta < \zeta^*$, then the choice is between ownership and provision by G, and ownership and provision by A. Comparing z_G^G and z_A^A, it is straightforward to show that K is larger under G authority if and only if

$$\rho = \frac{1 - \Delta\beta\alpha^A}{1 + \Delta\beta\alpha^G} < \frac{\theta^A}{\theta^G}.$$

Since $\rho < 1$, it is possible that in order to maximize the size of the facility, G should be given planning and expenditure authority, even if it has higher construction costs.

When asset transfer is not so costly, that is, $\delta > \zeta^*$, G should own the asset, and the choice is between G provision and A provision. By comparing z_G^G and z_A^G, it can be shown that K is larger when chosen by G if and only if

$$\rho' = \frac{1 - \Delta\beta\delta\alpha^G}{1 + \Delta\beta\alpha^G} < \frac{\theta^A}{\theta^G}.$$

This inequality indicates that, conditional on δ being large enough that G should own the facility (i.e., $\delta > \zeta^*$), as the cost of transferal increases it becomes more likely that the identity of the provider should differ from that of the owner. This somewhat counterintuitive result can be understood by examining the nature of the ex post bargaining game. As δ falls, the gains from ex post trade increase, because the surplus generated cooperatively is not affected by δ, but the non-cooperative surplus is reduced by a fraction $(1 - \delta)$. Now the net payoff to A is equal to its outside option $(\delta\beta_A W^G)$ which falls by $(1 - \delta)\beta_A W^G$, plus half of the gains from trade $(\frac{1}{2}(W^* - \delta W^G))$ which increase by $\frac{1}{2}(1 - \delta)W^G$. Thus A's payoff changes by

$$
\begin{aligned}
-(1-\delta)\beta_A W^G + \frac{1}{2}(1-\delta)W^G &= (\frac{1}{2} - \beta_A)(1-\delta)W^G \\
&= \frac{1}{2}\Delta\beta(1-\delta)W^G \\
&> 0.
\end{aligned}
$$

This higher payoff increases A's incentive to provide the capital good, so, ceteris paribus, it becomes more likely that A's choice will be higher than G's. Again, the public good nature of the output produced by the two parties is of importance for this result.[14]

Finally, the effect of an increase in α^G is also to increase the likelihood that G should provide the facility. In this case, the benefits of cooperation are relatively small, and the increase in G's outside option is relatively large compared with the increase (if any) in the gains from trade, thus increasing the private return to G from investing in K.

6 Concluding remarks

This paper has attempted to delineate a number of dimensions of decentralization and to formalize notions of participation and ownership in the design of health delivery systems. It is motivated in part by the simple observation that many different organizational forms for the delivery of health care exist, while their analysis is often descriptive at best. I have employed the well-established incomplete contracts paradigm of Hart and others to examine the organization of health delivery when the inputs of two parties are required, but when they are unable to write down, and commit to, a binding and complete contract describing the actions each should take.

The allocation of ownership of a health care facility (e.g., a hospital) to the central government or local authority affects each party's choice of ex interim effort - i.e., effort that is required to ensure efficient functioning of the facility. The strong result that emerges is that in order to maximize the joint surplus of the two parties (which, in this model, is equivalent to the welfare of local residents), ownership of the facility should be given to the party that puts the higher value on the welfare of local residents. This result is more robust that the corresponding results from the standard theory

---

[14] It should be noted again that, in equilibrium, the asset will be used cooperatively, and the surplus generated will not be reduced because of the transfer of ownership. Thus it is appropriate to compare levels of capital investment between ownership structures directly, and not to adjust for the productivity of capital when transferred (off the equilibrium path).

of the firm model because of the public good nature of the output that is produced. In particular, if ex post bargaining between the parties breaks down, each still enjoys some benefit from the other's actions.

The second result is that in general it is optimal to allocate tax authority to the other party, the one that values the well-being of residents less. That is, expenditure responsibility (ownership) and financing authority should be separated. The intuition for this is that a higher financing burden (induced by access to a small and inefficient tax base) has the same incentive effect on effort as does asset ownership - both increase the return to effort.

Finally, the paper examines the situations under which inefficient production of the underlying facility at the ex ante stage is optimal. That is, can it be desirable to induce "participation" by otherwise inefficient providers? Even if it is costless to transfer ownership of the health facility from the provider to the other party, it may still be true that the party with higher costs should have authority over the choice of facility and retain ownership, instead of "contracting out" this decision. Somewhat paradoxically, within a range of relatively low transfer costs, the more costly it is to transfer the asset from the provider to the owner (when they differ), the more likely it is that ownership and provision should be separated.

7 References

Banerjee, Abijit (1997): "A Theory of Misgovernance," *Quarterly Journal of Economics*, 1289-

Bardhan, Pranab and Dilip Mookherjee (1999): "Relative capture of local and central governments: An essay in the political economy of decentralization," mimeo, department of economics, University of California, Berkeley.

Besley, Timothy and Maitreesh Ghatak (1999): "Public-Private Partnerships for the Provision of Public Goods: Theory and an Application to NGOs in Developing Countries," presented at a World Bank seminar, August.

24

Filmer, Deon, Jeffrey Hammer, and Lant Pritchett (1998): "Health Policy in Poor Countries: Weak Links in the Chain," Policy Research Working Paper # 1874, World Bank.

Grossman, Sanford and Oliver Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, **94**(4), 691-719.

Hart, Oliver (1995): *Firms, Contracts, and Financial Structure*, Oxford University Press.

Hart, Oliver, Andrei Shleifer, and Robert Vishny (1998): "The Proper Scope of Government: Theory and an Application to Prisons," *Quarterly Journal of Economics*, **112**(4), 1119-1158.

Holmstrom, Bengt and Paul Milgrom (1987): "Aggregation and Linearity in the provision of Intertemporal Incentives," *Econometrica*, 55(2), 303-328.

Holmstrom, Bengt and Paul Milgrom (1991): "Multi-task Agency Problems," *Journal of Law, Economics, and Organization*

Mills, Anne (1994): "Decentralization and accountability in the health sector from an international perspective: what are the choices?" *Public Administration and Development*, 14, 281-292.

Oates, Wallace (1999): "An Essay on Fiscal Federalism," *Journal of Economic Literature*, XXXVII(3), 1120-45.

Seabright, Paul (1996): "Accountability and Decentralisation in Government: An Incomplete Contracts Model," *European Economic Review*, 40, 61-89.

World Bank (1993): *World Development Report 1993: Investing in Health*, Oxford University Press.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2378	Disintegration and Trade Flows: Evidence from the Former Soviet Union	Simeon Djankov Caroline Freund	June 2000	R. Vo 33722
WPS2379	India and the Multilateral Trading System after Seattle: Toward a Proactive Role	Aaditya Mattoo Arvind Subramanian	June 2000	L. Tabada 36896
WPS2380	Trade Policies for Electronic Commerce	Aaditya Mattoo Ludger Schuknecht	June 2000	L. Tabada 36896
WPS2381	Savings and the Terms of Trade under Borrowing Constraints	Pierre-Richard Agénor Joshua Aizenman	June 2000	T. Loftus 36317
WPS2382	Impediments to the Development and Efficiency of Financial Intermediation in Brazil	Thorsten Beck	June 2000	E. Mekhova 85984
WPS2383	New Firm Formation and Industry Growth: Does Having a Market- or Bank-Based System Matter?	Thorsten Beck Ross Levine	June 2000	E. Mekhova 85984
WPS2384	Are Cost Models Useful for Telecoms Regulators in Developing Countries?	Daniel A. Benitez Antonio Estache D. Mark Kennet Christian A. Ruzzier	July 2000	G. Chenet-Smith 36370
WPS2385	The Rise, the Fall, and ... the Emerging Recovery of Project Finance in Transport	Antonio Estache John Strong	July 2000	G. Chenet-Smith 36370
WPS2386	Regulators and the Poor: Lessons from the United Kingdom	Richard Green	July 2000	G. Chenet-Smith 36370
WPS2387	The Long and Winding Path to Private Financing and Regulation of Toll Roads	Antonio Estache Manuel Romero John Strong	July 2000	G. Chenet-Smith 36370
WPS2388	The Role of Special and Differential Treatment for Developing Countries in GATT and the World Trade Organization	Constantine Michalopoulos	July 2000	L. Tabada 36896
WPS2389	Vietnam: On the Road to Labor-Intensive Growth?	Patrick Belser	July 2000	H. Sutrisna 88032
WPS 2390	The Social Rate of Return on Infrastructure Investments	David Canning Esra Bennathan	July 2000	H. Sladovich 37698

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2391	Are the Poor Protected from Budget Cuts? Theory and Evidence for Argentina	Martin Ravallion	July 2000	P. Sader 33902
WPS2392	What Factors Appear to Drive Private Capital Flows to Developing Countries? And How Does Official Lending Respond?	Dipak Dasgupta Dilip Ratha	July 2000	S. Crow 30763
WPS2393	Will the Euro Trigger More Monetary Unions in Africa?	Patrick Honohan Philip R. Lane	July 2000	A. Yaptenco 31823
WPS2394	Tax Evasion, Corruption, and the Remuneration of Heterogeneous Inspectors	Waly Wane	July 2000	H. Sladovich 37658