Corruption and Decentralization of Infrastructure Delivery in Developing Countries

Pranab Bardhan

University of California, Berkeley

Dilip Mookherjee

Boston University

June 2001

Corresponding Author: Dilip Mookherjee, Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215. Tel. 617-353-4392; Fax: 617-353-4143; email: dilipm@bu.edu.

This paper extends the analysis in Part 2 of our earlier working paper titled 'Expenditure Decentralization and Delivery of Public Services in Developing Countries". For financial support we are grateful to the MacArthur Foundation. Mookherjee's research was additionally funded by National Science Foundation Grant No SBS-9709254. For helpful comments we are grateful to participants at the 1998 meeting of the MacArthur Research Network on Inequality and Economic Performance, and conferences at UC Irvine in February 2001 and at Yale in March 2001.

Abstract

Corruption and targeting failures in the delivery of public services in developing countries has frequently been argued to result from absence of controls on the behavior of central bureaucrats delegated authority over service delivery. This has motivated recent initiatives towards decentralization of service procurement and delivery to elected local governments expected to be more accountable to user interests. However if local democracy is prone to capture by local elites, decentralization can also be subject to diversion and targeting failures. This paper presents a simple analytical framework to evaluate the resulting trade-off, and predict the effects of decentralization on volume and allocation of service delivery under different financing mechanisms. Consistent with existing cross-country empirical evidence, we find that greater fiscal autonomy of local governments expands the volume of service delivery, but this tends to be accompanied by service overprovision to local elites at the expense of non-elites. Restrictions on the ability of local governments to raise local taxes can accordingly be justified on efficiency and equity grounds. User fee mechanisms ensure that decentralization welfare dominates centralization, irrespective of the degree of local capture.

Keywords: bureaucracy, corruption, decentralization, local governments, infrastructure, targeting; JEL Classification Nos.: D72, D73, H 41, H42, H77, I38, O17

1 Introduction

Decentralization of public service provision to local governments is a worldwide trend, and especially in developing countries (World Bank (2000)). A major motivating factor is the perceived lack of accountability of traditional centralized service delivery mechanisms. In particular it is widely believed that weaknesses in information and control systems in developing country bureaucracies have traditionally given rise to significant cost-padding, service diversions, and corruption among

centrally appointed bureaucrats.[1] Owing to weak supervision and communication systems, elected central government officials lack information about costs and needs of various local communities, and are unable to monitor actual service delivery patterns, thus permitting substantial abuse of power by the bureaucrats they entrust delivery to. Accordingly devolution of power over service procurement and delivery to elected local governments represents a possible way of enhancing accountability in the delivery mechanism. Local users can evaluate local cost and need, and monitor actual delivery patterns; hence nonperforming local administrations can expect to be ejected from office in the presence of a functioning local democracy.[2] This perspective has motivated decentralization and 'community targeting' mechanisms worldwide, such as of safety net programs in Albania (Alderman (1998)), education programs in Bangladesh (Galasso and Ravallion (1999)), participative budgeting in Porto Alegre, Brazil (Santos (1997)), neighborhood committees in China (Wong (1994)), the 73rd and 74th Constitutional Amendments in India creating a nationwide third tier of elected local governments in India (Isaac and Franke (2000)), schooling reform in El Salvador (Jimenez and Sawada (1999)), fiscal decentralization in South Africa in the post-apartheid regime (World Bank (2000)), and child benefit programs in Uzbekistan (Coudouel, Marnie et. al. (1998), to name only a few examples.

This enthusiasm has however been tempered by the awareness that local democracy may not function effectively in many developing countries. In particular, they may be subject to capture by local

[1] Analysis, examples and empirical evidence concerning such 'leakages' and 'targeting failures' are provided by Banerjee (1997), Bardhan (1996a, 1996b), Besley (1989), Besley and Kanbur (1993), Bird (1995), Dreze and Saran (1995), Grosh (1991, 1995), Lipton and Ravallion (1995), van de Walle and Nead (1995), and the 1990, 1994 and 1997 World Development Reports.

[2] Additional benefits of 'yardstick' competition and citizens 'voting with their feet' when services are delivered by diverse local governments may also enhance accountability. The significance of such Tiebout-style mechanisms in the context of developing countries is however open to debate, given high mobility and information costs across communities. Accordingly we shall not emphasize these in our analysis.

elites, to an extent depending on traditions of political participation, voter awareness and literacy, allocation of social and economic power within communities, lobbying and campaign contributions by wealthier groups, fairness and regularity of elections, transparency in local decision-making processes and government accounts, media attention and so on. These can vary widely across communities and countries, as documented in numerous studies by political scientists (e.g., Crook and Manor (1998)) and case studies of effects of decentralization (e.g., Conning and Kevane (1999) Isaac and Franke (2000), Lieten (1996), Mathew and Nayak (1996) and World Bank (2000)). Thus decentralization is also fraught with the danger of significant targeting failures, as local elites seek to divert services to themselves at the expense of non-elites. It is not clear on *a priori* grounds whether and when decentralization will ensure genuine improvements in efficiency and equity in service deliveries.

The available empirical evidence pertains mainly to cross-country comparisons. In a study covering 57 countries over the period 1980-95, Fisman and Gatti (1999) document a significant negative correlation between the subnational share of total government spending and various measures of corruption.[3] Estache and Sinha (1995) study 20 countries over the period 1970-92 and find a significant positive effect of expenditure decentralization on per capita infrastructure deliveries. They also find that the effect is stronger in developing countries compared with developed countries, and is weaker when local governments rely more on central funds rather than their own revenues. Other forms of evidence include case studies, of which the comparative study of Korean and Indian irrigation bureaucracies by Wade (1997) is particularly notable. This study pinpointed the overcentralized top-down system in India accompanied by weak communication and monitoring systems as the principal cause of poorer delivery performance relative to Korea.

[3]They control for level of per capita income, size of the country, index of civil liberties, openness to trade, and ethnic fractionalization. Their results are also robust with respect to possible endogeneity bias: the negative correlation survives when legal origin is used as an instrument for decentralization.

This paper develops a simple analytical framework that formalizes the tradeoff between centralized and decentralized delivery systems, in the context of infrastructure services such as roads, water, electricity or telecommunications.[4] To represent the principal effects as starkly as possible, we consider stylized polar versions of either system. In the centralized system we suppose that authority is assigned entirely to bureaucrats whose objective is to maximize their net incomes, i.e., bribes less the cost of delivery, so behave like unregulated monopolists. Decentralization shifts control rights to a local government which seeks to maximize a weighted sum of welfares of two classes of local users: elites and non-elites, where elites value the service more. Elites receive a higher welfare weight, representing the phenomenon of capture of local governments: we impose no particular restriction on the extent of such capture.

Three different financing mechanisms for local governments are considered, corresponding to those principally observed in different countries (Dillinger (1995)): (i) local taxes, (ii) user fees, and (iii) central grants. The first two forces local governments to be entirely self-financed. In practice the capacity and authority to finance themselves from local revenues of many local governments is restricted considerably, either due to limited collection and enforcement abilities, or to constitutional restrictions on their ability to levy local taxes and fees. In such circumstances, local governments rely substantially on fiscal grants from the central government. The mode of financing local governments is accordingly a key factor in the institutional design of a decentralized system.

Our principal results are the following. Owing to the monopoly power of bureaucrats (combined

[4]Bardhan and Mookherjee (2000b) explore similar issues in the context of delivery of an anti-poverty program. The contrast with the case of an antipoverty program is that the latter is targeted exclusively to poor citizens, who lack the purchasing power necessary to purchase the service either via bribes or user fees. Neither can they shoulder any tax burden. Accordingly the analysis of both centralized and decentralized regimes differs considerably from the case of infrastructural services.

with restrictions on their ability to engage in perfect 'bribe'-discrimination), the centralized system is subject to inefficiently low and inequitable patterns of service deliveries. Decentralization typically tends to expand service deliveries, as authority shifts to those more responsive to user needs. However in the presence of capture of local governments, there is a tendency for the service to be overprovided to local elites, at the expense of non-elites. The extent of such cross-subsidization — which is both inefficient and inequitable — depends on the extent of local capture, and on the degree of fiscal autonomy of local governments. With maximal autonomy, i.e., local tax finance, the welfare comparison with centralization depends on the degree of capture: centralization will be superior if (and only if) there is sufficient local capture.

With user fee financing, however, this can no longer happen: decentralization unambiguously welfare dominates centralization, as well as local-tax-financed-decentralization, *irrespective of the extent of local capture*. User fees force local governments to rely on noncoercive mechanisms, in which the participation of local users is voluntary: this restriction on the capacity of local governments to expropriate non-elites prevents excessive cross-subsidization in favor of elites. At the same time it allows flexible adjustment of service deliveries to changes in local need or cost.

In contrast, central grant financing is less flexible, owing to the tendency of local governments to claim higher local need or cost *vis-a-vis* other communities and the central government. This restricts the level of service deliveries relative to user-fee-financing, by an extent that depends on a range of additional parameters (such as the deadweight costs of central revenues, relative capture of central and local governments, and local need and cost shocks). The extent of service overprovision to elites is also correspondingly restricted. The welfare implications (relative to centralization or user fees) are therefore ambiguous, and depend on a range of relevant political and financing parameters.

The most important result thus concerns the welfare superiority of decentralization financed

by user fees, which applies no matter how poorly local democracy functions. This suggests that decentralization initiatives suitably designed to restrict the fiscal capacity of local governments can help prevent the worst evils of either system. Under the other financing mechanisms, decentralization may or may not generate a welfare loss relative to centralization, depending on a host of relevant parameters. Our model also generates predictions concerning the volume of service deliveries across different regimes, which are consistent with the cross-country empirical evidence. At the same time the model cautions against ranking these regimes on the basis of aggregate service volumes alone, since these conceal targeting misallocations and the possibility of service overprovision to elites under decentralization. It also cautions against the use of corruption measures based on bribes, as they exclude forms of political corruption resulting under decentralization.

The issue of corruption and lack of accountability has largely been ignored by previous literature on fiscal decentralization. The traditional literature on fiscal federalism (e.g., summarized in Cremer, Estache and Seabright (1995), Oates (1972), Musgrave and Musgrave (1984, Chapter 24) and Inman and Rubinfeld (1996, 1997)) identifies uniformity and standardization as the key drawback of centralized provision, and interjurisdictional spillovers and inability to exploit scale economies as the main drawback of decentralization. A number of authors have criticized this literature for not providing a compelling or realistic account of the drawbacks of centralized provision, and have relied instead on political economy considerations (e.g., Besley and Coate (1999), Bolton and Roland (1997), Laffont and Pouyet (2000), Lockwood (1998), Seabright (1996) and Tommasi and Weinschelbaum (1999)). Among them, only Seabright (1996) and Tommasi and Weinschelbaum (1999) focus on lack of accountability as the principal drawback of centralization, which has to be traded off against interjurisdictional coordination problems inherent in decentralization. Our theory in contrast is based on the view that centralized systems may be more or less accountable than local governments, depending on the nature of political institutions. The basis for this view has been argued by

Bardhan and Mookherjee (1999, 2000a). In particular, our approach accommodates the frequently expressed concern in many developing countries that high levels of capture of local governments may worsen intracommunity misallocations under decentralization, compared with centralization. At the same time our model focuses less on problems of interjurisdictional coordination.[5]

Section 2 introduces the model, Section 3 describes the centralized regime, and Section 4 the various decentralization regimes. Section 5 concludes.

2 The Model

The service is a private benefit such as irrigation or electricity, whose production is subject to a large fixed cost F, in addition to variable costs. Production is concentrated in a single large utility. There are a number of different communities, denoted $i = 1, \ldots, n$. The variable cost of generating supply Y_i to community i is $\theta_i Y_i$, so θ_i is the (constant) marginal cost of delivery to community i. The realization of θ_i is random, represented by a positive density function t_i over the interval $[\underline{\theta}_i, \bar{\theta}_i]$.

Community i has N_i users, who belong to either of two groups, large (l) and small (s), who differ in their valuation of the service. A fraction β_i of citizens in the community are small users; the rest are large users. Large users belong to a wealthy elite, and value the service more. The utility function of a member of group $k = l, s$ in community i is $\gamma_k \eta_i v(y_k) - t_k$, where γ_k is a group-specific

[5]For instance there are no capacity constraints on service delivery across different communities. We also abstract from the possibility that local governments may possess less administrative or technical competence relative to central bureaucrats. Our model does however accommodate economies of scale in service provision across communities, which motivates production to be concentrated in a single utility from which local governments procure the service. It also allows for the tendency for local communities to free-ride off revenues raised by the central government in a system where local governments are financed by fiscal grants from the central government.

valuation parameter satisfying $\gamma_l > \gamma_s > 0$, η_i is a community-specific need shock, y_k is the level of service delivered, and t_k is the net financial burden imposed on the user. The utility function v is homothetic: $v(y) = \frac{1}{\alpha+1} y^{\alpha+1}$ where $\alpha < 0$ and different from -1. Local need η_i is distributed independently across regions; within region i it has a positive density function h_i on an interval $[0, \eta_u]$ which satisfies a standard monotone hazard rate condition that $\frac{1-H_i}{h_i}$ is nonincreasing, where H_i denotes the corresponding distribution function.

The central government knows only the demographic profile of the different communities, i.e., the populations N_i and its composition among the two groups β_i. It does not observe the realization of local need or cost shocks, and cannot monitor local service delivery patterns. Central taxes involve a deadweight cost of $\lambda > 0$. Taking these deadweight costs into account, the second-best service allocation is given by solution to

$$\gamma_k \eta_i v'(y^f_{ki}) - (1+\lambda)\theta_i = 0. \tag{1}$$

The corresponding *first-best allocation* corresponds to the case where the deadweight costs of finance λ equals zero.

3 Centralized Bureaucracy

Under centralization, authority over service delivery is delegated to bureaucrats appointed by the central government. The bureaucracy consists of two layers. The top layer (which may include central government officials upto the highest level) is in charge of the central utility and allocates services across communities, i.e., decides Y_i. The bottom layer consists of bureaucrats assigned to local communities, who allocate the community service allocation Y_i across different local users. Top layer bureaucrats observe the realization of cost θ_i for each community i, but not the local need η_i.

The lower level bureaucrat assigned to the community observes the realization of η_i. However, being outsiders assigned temporarily to any given community, they cannot distinguish different categories of users within it. Hence attempts to charge higher bribes to the large users that value the service more can be circumvented by these users by masquerading as a collection of small users (e.g., by splitting their lands and assets among different family members).

Consider first how a local bureaucrat would optimally 'sell' a given level of aggregate service Y_i to region i among different local users, when the nature of local need and delivery cost happens to be (η_i, θ_i). Bribe and service levels for the two classes (denoted by b_k and y_k respectively) would be set to maximize (per capita) bribe income $\beta_i b_s + (1-\beta_i)b_l$, subject to voluntary participation constraint for each class k: $\gamma_k \eta_i v(y_k) - b_k \geq 0$, the incentive constraint that large users do not seek to masquerade as small users[6]: $\gamma_l \eta_i v(y_l) - b_l \geq \gamma_l \eta_i v(y_s) - b_s$, and the allocation constraint $\beta_i y_s + (1-\beta_i) y_l \leq \frac{Y_i}{N_i}$. Standard arguments can be employed to show that the participation constraint binds for small users, and so does the incentive constraint for large users, implying that

$$b_s = \gamma_s \eta_i v(y_s), b_l = \gamma_l \eta_i v(y_l) - (\gamma_l - \gamma_s)\eta_i v(y_s). \tag{2}$$

This generates the following reduced form expression for bribe income as a function of service delivery levels:

$$\beta_i D_s \eta_i v(y_s) + (1-\beta_i) D_l \eta_i v(y_l) \tag{3}$$

where $D_s \equiv \gamma_s - \frac{1-\beta_i}{\beta_i}(\gamma_l - \gamma_s)$ and $D_l \equiv \gamma_l$ represent the 'virtual' valuation parameters for the two classes respectively. Maximizing bribe income (3) less variable delivery cost yields (7). Moreover,

[6] It is well known that the solution will automatically satisfy the reverse incentive constraint as well, so small users will not try to masquerade as large users either. Moreover, the solution will automatically have the property that no large user will seek to masquerade as m (> 1) small users, since the large users will be charged a discounted bribe rate relative to small users.

the local bureaucrat ends up with a total bribe income from this community of

$$N_i B_i \eta_i \frac{Y_i^{-\frac{1}{\alpha}}}{\alpha + 1} \tag{4}$$

where $B_i \equiv [\beta_i D_s^{-\frac{1}{\alpha}} + (1 - \beta_i) D_l^{-\frac{1}{\alpha}}]^{-\alpha} < 1$.

Next turn to the allocation of service levels by higher level bureaucrats. These bureaucrats will seek to extract bribe kickbacks from lower level bureaucrats in exchange for allocating service levels to their respective communities. The extent to which they can do so will be restricted by the superior information of local bureaucrats concerning local need η_i, which defines the bribe income that can be earned in community i from a given service allocation Y_i, as given by (4). Applying the Revelation Principle, without loss of generality they design a mechanism for the local bureaucrat in community i, given by $Q_i(\eta_i), Y_i(\eta_i)$, defining the kickback Q_i that the local bureaucrat must pay to the central bureaucrat, and the corresponding service level allocated to the local bureaucrat, consequent on a reported valuation parameter η_i by the latter. The central bureaucrat selects these mechanisms, one for each local bureaucrat, to maximize their surplus, i.e., the difference between aggregate kickback and production costs

$$\mathcal{E} \sum_i [Q_i(\eta_i) - \theta_i Y_i(\eta_i)] \tag{5}$$

whereas the local bureaucrat is motivated to maximize the difference between local bribe income and the kickback that needs to be paid to their bosses.[7] Hence the maximization (5) is subject to a breakeven and truthful reporting constraints for each local bureaucrat:

$$N_i B_i \eta_i v(Y_i(\eta_i)) - Q_i(\eta_i) \geq 0$$

[7]Owing to the inability of the central government officials to monitor service deliveries, the budgetary support C provided by the government to the utility is a lumpsum amount. High level bureaucrats seek to maximize the expected value of sum of budgetary slack $C - F - \sum_i \theta_i Y_i$ and bribe kickbacks $\sum_i Q_i$, from which the objective function (5) follows.

$$\eta_i \in \arg\max_{\tilde{\eta}_i}[N_i B_i \eta_i v(Y_i(\tilde{\eta}_i)) - Q_i(\tilde{\eta}_i)]$$

Again standard techniques of solving these principal-agent problems (e.g., based on Baron-Myerson (1982)) can be employed to show that the equilibrium intercommunity service allocation $Y_i(\eta_i, \theta_i)$ maximizes

$$\sum_i [N_i B_i J_i(\eta_i) v(Y_i) - \theta_i Y_i]. \tag{6}$$

where $J_i(\eta_i)$ denotes $\eta_i - \frac{1-H_i(\eta_i)}{h_i(\eta_i)}$.

We thus obtain the following outcome under centralization.

Proposition 1 *The centralized system results in the following allocation (for any given state of the world* $(\theta_i, \eta_i), i = 1, 2, \ldots$.

(i) In any given state $\{(\theta_i, \eta_i)\}_{i=1,\ldots,m}$*, the community service allocation* $Y_i(\eta_i, \theta_i)$ *maximizes (6), and results in underprovision relative to the first-best.*

(ii) Intracommunity allocation (given the per capita service level Y_i*) for community* i *is given by*

$$\begin{aligned} y_s^* &= Y_i \frac{D_s^{-\frac{1}{\alpha}}}{\beta_i D_s^{-\frac{1}{\alpha}} + (1-\beta_i) D_l^{-\frac{1}{\alpha}}} \\ y_l^* &= Y_i \frac{D_l^{-\frac{1}{\alpha}}}{\beta_i D_s^{-\frac{1}{\alpha}} + (1-\beta_i) D_l^{-\frac{1}{\alpha}}} \end{aligned} \tag{7}$$

and results in further underprovision to low valuation (i.e., small users relative to large users). Small users obtain a net utility of 0*, while large users obtain a positive surplus, with the bribes given by (2).*

Competition for rents across different layers of the bureaucracy causes the intercommunity allocation to be skewed in favor of communities with high need. Lower level bureaucrats are tempted to

understate the η_i for their community in order to limit the kickback they have to pay their superiors. This temptation is counteracted by underproviding the service to a community when a low η_i is reported. This distortion compounds the distortion resulting from inability of local bureaucrats to price discriminate perfectly (manifested by $B < 1$). The end result is (i) underprovision of service levels to each community (relative to the first-best allocation that corresponds to a zero deadweight cost of taxes).[8], and (ii) the intracommunity distortion whereby service delivery is underprovided to small farmers.

4 Decentralization

Now suppose authority over service delivery is devolved to local governments. They procure the service from the central utility, and allocate it across local users. In order to focus on considerations related to local capture (rather than the possibility of limited technical or administrative competence of local government officials that might raise costs under decentralization), we assume that they know the delivery cost θ_i and procure Y_i at this cost from the central utility.[9]

The local government may be captured by local elites owing to a variety of distortions in the functioning of local democracy. We represent the objective of the local government in community i by

$$W_i^l = \beta_i U_{si} + \delta_i^l (1 - \beta_i) U_{li} \tag{8}$$

where $\delta_i^l > 1$ represents the premium placed on the welfare of elites relative to non-elites. The switch from centralization to decentralization thus shifts control rights away from bribe extractors

[8] With a positive deadweight cost, whether there is under or over-provision depends on how large λ is.

[9] In the absence of this assumption, decentralization will be subject to an additional disadvantage relative to centralization.

to those who do respond to the interests of local users, owing to the operation of electoral pressures. However, they respond with a bias in favor of local elites, and in extreme situations are effectively owned and operated by these elites. This bias may reflect the preeminence of wealth and social status of elites within the community, their connections with the outside world, political awareness and involvement in local politics relative to non-elites, ability to contribute to campaign funds of political parties, control local elections, media, or force. Nevertheless some degree of responsiveness of local governments to the interests of small users arises from the fact that these users vote in local elections: any local government that rides roughshod over their interests may be ejected from office by disgruntled voters. Accordingly, the degree of capture may depend on β_i, the demographic weight of small users within the community. We impose no particular structure on the capture coefficient δ_i^l, as it summarizes a multitude of political determinants of local capture that we take to be exogenous.

One particular model of electoral competition that generates an objective function exactly of the form (8) is the Baron (1994) or Grossman-Helpman (1996) theory of special interest groups that contribute to campaign finance of two parties or candidates engaging in Downsian competition for local office. This version is elaborated further in Bardhan and Mookherjee (1999, 2000a). In that version it turns out that the extent of capture δ_i^l is an increasing function of β_i, owing to the lower level of political awareness among small users, which increases the value of campaign funds in winning elections (thus increasing the influence of elites arising from their campaign contributions). Moreover, δ_i^l tends to 0 as β_i tends to 0 and to a finite limit as β_i tends to 1. While these assumptions are inessential to our results, the figures illustrating the service deliveries under different regimes will be drawn corresponding to such a case.

We consider alternative financing mechanisms for the decentralized regime.

4.1 Local-Tax-Financed Decentralization

In this version expenditure decentralization is accompanied by devolution of local revenue raising authority to local governments, which are fiscally autonomous and self-sufficient. So we assume that local governments have the ability and constitutional authority to finance their expenditure needs entirely from local revenues, at the same deadweight cost λ as the central government.

A local government will set service levels and local taxes y_k, t_k for the two classes $k = p, l$ to maximize

$$\beta_i[\gamma_s \eta_i v(y_s) - t_s] + \delta_i^l (1 - \beta_i)[\gamma_l \eta_i v(y_l) - t_l] \tag{9}$$

subject to the budget constraint $\beta_i t_s + (1 - \beta_i) t_l = (1 + \lambda)\theta_i[\beta_i y_s + (1 - \beta_i) y_l]$, and nonnegativity constraints on t_l, t_s.[10] Voluntary participation constraints do not need to be imposed, as local governments have the ability to impose coercive taxes, and citizens cannot move across districts owing to high mobility costs. The resulting outcome is described below, along with the welfare comparison with centralization.[11]

[10] The nonnegativity constraints prevent elites from using the local fiscal mechanism to directly expropriate non-elites. Instead reverse redistribution must be carried out indirectly in the form of distorted patterns of service delivery and selective tax evasion by elites, rather than lumpsum transfers from non-elites to elites. Such forms of redistribution would typically be illegal. If these laws are not adequately enforced, the result would be to worsen equity under this mechanism, but would have no effect on the levels of service delivery or efficiency. Consequently this version of decentralization would perform worse relative to the other two variants (which would not be affected similarly), strengthening our results concerning ranking of different regimes.

[11] We use a utilitarian welfare criterion. However, the same results would apply with any other individualistic inequality-averse social welfare function, since the efficiency and equity comparisons across different regimes tend to be similar, i.e., the welfare comparisons do not require any efficiency-equity tradeoff.

Proposition 2 *Consider any state η_i such that*

$$1 + \lambda = \frac{\eta_i}{J(\eta_i)}. \tag{10}$$

Then local-tax-financed decentralization generates the following outcomes:

(i) tax burdens $t_l = 0$, $t_s = (1 + \lambda)\theta_i[y_s + \frac{1-\beta_i}{\beta_i} y_l]$, and service deliveries satisfying

$$\eta_i \gamma_s v'(y_s^d) = (1 + \lambda)\theta_i; \eta_i \gamma_l v'(y_l^d) = \frac{(1 + \lambda)\theta_i}{\delta_i^l} \tag{11}$$

i.e., second-best supply to small users and overprovision to large users, implying that service levels are larger than under centralization for both groups;

(ii) higher welfare than centralization as δ_i^l approaches 1;

(iii) lower welfare than centralization if δ_i^l is sufficiently large.

The role of assumption (10) is to abstract from a source of underprovision of per capita service delivery to the community in either system: the deadweight cost of local taxes in the decentralized system, and the competition for bribery rents across different layers of the bureaucracy in the centralized system. These two factors need not bear any relation to one another, thus introducing one set of parameters that affect relative service levels. Proposition 2 focuses instead on the pattern of intracommunity allocations, by controlling for these parameters.

The nature of the objective function (9) of the local government makes it evident that small users will bear the entire financial burden of the service: large users can use their political clout to evade their tax obligations entirely. Hence the objective of the local governments as a function of the service deliveries reduces to

$$\eta_i \gamma_s v'(y_s^d) = (1 + \lambda)\theta_i; \eta_i \gamma_l v'(y_l^d) = \frac{(1 + \lambda)\theta_i}{\delta_i^l} \tag{12}$$

from which result (i) follows. Decentralization thus results in a pattern of cross-subsidization representing reverse redistribution from non-elites to elites: the latter are overprovided the service at the expense of the non-elites. The greater the capture of local government the more extreme these misallocations, with lower efficiency and equity. On the other hand, with sufficiently low capture the local government maximizes welfare, so the allocation approaches the second-best. Hence the welfare comparison with centralization depends upon the extent of local capture.

Nevertheless, irrespective of the degree of capture, note that that our model predicts that decentralization expands the volume of infrastructural service delivered (assuming (10) holds). This is consistent with the empirical finding of Estache and Sinha (1995) in a cross-country context that expenditure decentralization results in increased supply of infrastructure services when accompanied by revenue decentralization. The principal reason for this in our model is the removal of monopoly (bribe) distortions inherent in the centralized system. Figure 1 depicts the nature of service allocations under the two systems across regions of varying demographic composition (corresponding to an increasing capture function $\delta_i^l(\beta_i)$ of the form predicted by the Grossman-Helpman model).

4.2 User Fee Financing

In practice, local governments in most developing countries lack elastic revenue bases, especially with respect to middle and low income citizens. Alternatively, local governments may lack the constitutional authority or administrative capacity to levy and collect local taxes. Local services correspondingly tend to be financed by fiscal grants from the center, whence local governments are no longer self-sufficient, creating a host of problems (such as asymmetric information about local need, 'soft' budget constraints, and dependence of service levels on the vagaries of public finances of the central government) that will be studied in the next section. One intermediate solution involves

local governments financing services by levying user charges, an approach commanding increasing attention in developing countries for infrastructure services. The virtues frequently commended for this approach are that they enhance the fiscal self-sufficiency of local governments. Less attention has been devoted to the implications for intracommunity allocations, to which we now turn.

The key feature of user fee financing (in contrast to local taxes) is their noncoercive character: fees are paid on the basis of voluntary purchase decisions by users. This has two important consequences. First, the government does not need a specialized administrative force to collect local taxes, thus limiting deadweight costs. Indeed, we shall assume that these are zero for collection of user fees: this is inessential to the arguments below, which will continue to apply as long as they do not exceed the deadweight costs of central tax revenues.[12]

Second, no local citizen can be made worse off as a result of the local government undertaking to deliver the service, no matter how much it is captured by local elites. Large users can of course still use their political power to evade payment of the nominal charges for the services they consume. Cross-subsidization therefore continues to be rendered possible by discriminatory user fees on the two categories of users. But the system must now respect voluntary participation constraints of small users, which limits the extent of such cross-subsidization. Formally, this involves adding the voluntary participation constraints to the optimization problem faced by the local government:

$$\eta_i \gamma_k v(y_k) - t_k \geq 0 \tag{13}$$

for both classes of users $k = s, l$.

[12]Note that the efficiency costs of user fees in terms of inducing over or under-use of the service are already incorporated in the analysis below, so the deadweight costs in the user fee mechanism involve only collection costs. These are likely to be much lower than administration and collection of direct taxes, which requires valuation of local properties and monitoring taxable activities of local citizens.

Proposition 3 *(i) Service and fees set by local governments under user-fee-financed decentralization are as follows:*

$$y_s = y_s^f, t_s = \gamma_s \eta_i v(y_s^f), y_l = \max\{y_l^f, \hat{y}_l\}, t_l = \theta_i(y_l^f - \hat{y}_l) \tag{14}$$

where $\hat{y}_l$ *denotes* $\frac{\beta_i(\gamma_s \eta_i v(y_s^f) - \theta_i y_s^f)}{(1-\beta_i)\theta_i}$. *Compared with centralization, service deliveries are larger for both groups. Compared with local-tax-financed decentralization, service deliveries are higher for small users, while the comparison is ambiguous for large users.*

(ii) User-fee-financed decentralization (weakly) Pareto dominates centralization: small users are equally well off while large users are better off. It welfare-dominates local-tax-financed decentralization, i.e., with respect to both efficiency and equity.

The reasoning is as follows. Consider the problem of maximizing the local government objective function (9) subject to (13) and the budget constraint $\beta_i t_s + (1-\beta_i)t_l = (1+\lambda)\theta_i[\beta_i y_s + (1-\beta_i)y_l]$. Clearly, the fee t_s for small users will be set at a level which reduces their surplus to zero, while providing them the first-best service level. The financial surplus generated thereby will be used to fund provision of the service to the large users. It pays for a service level y_l^f for large users. If this is larger than the first-best level y_l^f, this will be the service delivered to them. Otherwise the latter will pay the additional amount necessary to raise provision to the first-best level.

To prove (ii), consider first the welfare comparison with centralization. Small users are exactly as well off, since in either system they receive zero surplus. And large users are better off: this is obvious when $\hat{y}_l \geq y_l^f$, since they receive a larger service and pay nothing. In the other case they receive the first-best service level, the same as in centralization, and they pay less (they pay less than the cost of their service in the decentralized system, being subsidized partly by the small users, whereas they pay more than what their service costs at the margin under centralization in order to

generate positive rents for bureaucrats).

Next consider the comparison with decentralization financed by local taxes. Focus initially on the case where collection of user fees involves the same deadweight cost λ as local taxes. Then service levels under the user-fee system would shrink for both groups, and small users would get the same (second-best) service level under both systems. We claim that large users will be provided the service to a greater degree under tax financing. For if the solution to the local government's problem under tax financing generates nonnegative surplus for small users, the participation constraints for both user categories are automatically satisfied, the service levels must be the same. If it violates the participation constraint for small users, it must involve a larger payment by small users (since they receive the same service level under both systems). This must fund a larger service to large users under tax-financing. Since they receive second-best supply or greater under user-fee financing, the service will be (uniformly) over-provided to them under tax-financing. Therefore tax-financing must be both less efficient and less equitable. To complete the argument, note that if the collection of user fees involves lower deadweight costs compared to local taxes, then the relative performance of the user-fee mechanism improves even further.

A user-fee system administered by a local government subject to local capture will thus also overprovide the service to the large users at the expense of the small users, but to a lesser degree compared to the case of local tax finance. Service levels under the scheme are described in Figure 2. Supply to small users expands uniformly from second-best to first-best because of the reduction in deadweight costs of collection. The same is true for large users in regions with negligible (β_i close to 0) small users: supply levels expand from the second-best to the first-best level. In such regions there is an expansion of service deliveries to both groups, compared with the case of tax-financing. For regions with higher fraction of small users, there can be over-provision to large users, but to

a lesser degree under user-fee financing. Hence aggregate service level to communities can shrink with a restriction on the revenue raising capacity of local governments. From a welfare standpoint, however, this is a blessing — it reflects a mitigation of the damaging efficiency and equity effects of local capture.

This explains the welfare ranking of user-fee-financed decentralization relative to local tax finance or centralization. The generality of this result is striking: it holds irrespective of the degree of local capture, the composition of the district, or the realization of local need and cost shocks. Of course there are a number of qualifications: it rests on some of the maintained assumptions of the model, such as absence of inter-regional spillovers or redistribution, and adequate capacity of local governments to ensure cost-effective procurement. The argument also relies on the assumption that there are only two classes of users; it would be interesting to explore its robustness with an arbitrary number of different classes. Nevertheless, the result illustrates a number of advantages of user fees: lower collection costs, and limited scope for discretionary cross-subsidization by captured local governments to favor local elites. User fees selected by local governments with purchase decisions subsequently decentralized to individual users permit flexibility of service provision with respect to local cost and need. It is not, however, an optimal mechanism, since it permits some degree of cross-subsidization and overprovision to large users. This motivates interest in other financing mechanisms that restrict discretion of local governments in different ways.

4.3 Central Grant Financing

Now turn to the third principal financing mode commonly observed in many countries: grants from the central government. Suppose that local governments have no revenue raising capabilities at all, and receive block grants from the center to fund service allocation. Note that the local government

would always prefer a larger grant to a smaller one. This will give rise to an incentive problem between central and local governments: the latter would always like to overstate local need and cost in order to be eligible for a larger grant. Lacking information about local conditions, and being unable to monitor service deliveries actually implemented by local governments, the center will be unable to verify the claims of local governments. Consequently grants will end up being insensitive to local conditions, except demographic characteristics; this is exactly how grant formulae tend to be structured in practice.

The insensitivity of central grants to local conditions implies that first-best or second-best allocations cannot be implemented, even if all other conditions were ideal (e.g., if local governments were not subject to capture at all) — since welfare optimal allocations necessitate expenditures that do depend on local conditions. Nevertheless, some flexibility is possible if the grants are not tied to specific categories of services: local governments can then allocate a given budget across different services as their relative local needs and costs vary. Even tied grants admit considerable *de facto* fungibility, allowing them to be spent on alternative services via creative accounting practices: e.g. funds earmarked for construction of roads or canals can be spent on constructing a wide category of facilities, many of which privately benefit local elites.[13]

To represent such flexibility in its simplest form, we assume that governments allocate their fiscal resources between the infrastructure service in question, and lump sum subsidies to either class of users (representing the pecuniary equivalent of alternative services delivered). Since there is scope

[13] While diversions are possible from funds received from the central government, our analysis of the two previous regimes assumed implicitly that they are not possible from funds raised locally i.e., from either local taxes or user fees (which would represent a form of elite expropriation of small users within the community). This reflects the common tendency for communities to tolerate diversions of government funds to elites as long as those funds originated outside the community. Relaxing this assumption would reduce the welfare associated with the other two regimes.

for discretion with respect to whom the alternative services are delivered, we assume that the local government can discriminate perfectly across the users in their allocation. The results below will not be qualitatively altered with alternative specifications, e.g., if the alternative services are allocated uniformly across both classes, or if marginal utility of users are diminishing with respect to the level of their supply.

Given a fixed (per capita) block grant G, the local government in region i will select an allocation of the given infrastructure service y_s, y_l and subsidy equivalents S_s, S_l for alternative services for the two classes of users to maximize

$$\beta_i[\gamma_s \eta_i v(y_s) + S_s] + \delta_i^l(1-\beta_i)[\gamma_l \eta_i v(y_l) + S_l] \tag{15}$$

subject to the budget constraint

$$\beta_i[\theta_i y_s + S_s] + (1-\beta_i)[\theta_i y_l + S_l] \leq G \tag{16}$$

and the nonnegativity constraints $S_s \geq 0, S_l \geq 0$ that arise from the lack of local revenue raising capacity. Given local capture it is immediately evident that small users will receive no alternative services at all: $S_s = 0$. Hence grant income not spent on the infrastructural service will be diverted to the procurement of alternative services that benefit local elites selectively.

Since the budget constraint (16) must bind, it follows that $S_l = G - \theta_i[\beta_i y_s + (1-\beta_i)y_l]$: the opportunity cost of spending more resources on the assigned service is that less is available for diversion to the alternative service. *In contrast to the two previous financing modes, therefore, the cost of service delivery at the margin is effectively borne by large rather than small users.* This causes the resulting service allocations to differ qualitatively. The problem of the local government reduces to maximization of

$$\eta_i[\beta_i \gamma_s v(y_s) + \delta_i^l(1-\beta_i)\gamma_l v(y_l)] + \delta_i^l\{G - \theta_i[\beta_i y_s + (1-\beta_i)y_l]\} \tag{17}$$

subject to the constraint that $G \geq \theta_i[\beta_i y_s + (1-\beta_i)y_l]$. If the grant G is large enough, this constraint will not bite, and the service allocations will satisfy the first order condition

$$\gamma_s \eta_i v'(y_s) = \delta_i^l \theta_i; \quad \gamma_l \eta_i v'(y_l) = \theta_i \tag{18}$$

Large users then get delivered the first-best level, while there is under-provision to small users. The pattern bears a closer similarity to that resulting under centralization, rather than the two other modes of decentralization considered previously.

The implications of a given block grant G on community allocation is described next.

Proposition 4 *The allocation resulting from a block grant G in community i is the following:*

(i) Service delivery for a group k user is $y_k = f_k.Y_i^l$, where Y_i^l is the per capita service level in the community (described further below), and f_k is the share of group k, determined as follows: $f_s = \frac{{\gamma_s}^{-\frac{1}{\alpha}}}{\beta_i {\gamma_s}^{-\frac{1}{\alpha}} + (1-\beta_i)\delta_i^l {\gamma_l}^{-\frac{1}{\alpha}}}$, *and* $\beta_i f_s + (1-\beta_i) f_l = 1$.

(ii) There exists a threshold need level η_i^* that depends on G, θ_i, δ_i^l[14] such that when local need θ_i is less than η_i^*, the local government is not financially constrained, with the per-capita service level for the community Y_i^l equal to the desired level $Y_i^f(\eta_i, \theta_i)$, characterized by

$$L_i \eta_i v'(Y_i^f) = \delta_i^l \theta_i \tag{19}$$

where L_i denotes $\frac{\beta_i {\gamma_s}^{1-\frac{1}{\alpha}} + (1-\beta_i)\delta_i^l {\gamma_l}^{1-\frac{1}{\alpha}}}{\beta_i {\gamma_s}^{-\frac{1}{\alpha}} + (1-\beta_i)\delta_i^l {\gamma_l}^{-\frac{1}{\alpha}}}$. *In this case spending on the service is less than the grant G, with the surplus diverted to elite consumption ($S_l > 0$). When need exceeds η_i^*, the local government is financially constrained, spending it entirely on the service, so $Y_i^l = \frac{G}{\theta_i} < Y_i^f$, and there is no diversion.*

[14] This is given by $\eta_i^*(G, \theta_i, \delta_i^l) = \frac{\delta_i^l \theta_i}{L_i v'(\frac{G}{\theta_i})}$.

The per capita service delivery pattern is thus

$$Y_i^L(\eta_i, \theta_i, G) = \min[Y_i^f(\eta_i, \theta_i), \frac{G}{\theta_i}] \tag{20}$$

as depicted in Figure 3. This restriction in the flexibility of service levels to local conditions in high need states is a distinctive feature of grant-financed decentralization, which owes to the lack of fiscal self-sufficiency of local governments, the nonverifiability of local needs and costs by the central governments, and the accompanying tendency for local governments to attempt to free-ride off the revenues raised from other communities. The severity of these fiscal constraints depends on how large the grant is. We therefore need to turn to the question of how these grants are determined.

This depends on the political objectives of the central government, and the way that it raises revenues. The central government may also be subject to capture by elites, to an extent that may bear no obvious relation to the extent of local capture (as argued in Bardhan and Mookherjee (1999, 2000a)). So letting δ^c denote the degree of capture at the central level, the objective of the central government is

$$\Sigma_i N_i[\beta_i U_{si} + \delta^c(1-\beta_i)U_{li}] \tag{21}$$

As for financing patterns, it is well known that for a variety of reasons, both including political will and administrative ease, most developing countries rely primarily on indirect (sales, excise and customs duties) rather than direct taxes (see, e.g., the evidence cited in Ahmed and Stern (1984), Newbery and Stern (1987) and Das-Gupta and Mookherjee (1998)). Owing to their regressive nature, we shall assume that small users bear a burden that is proportionately greater or the same as the burden borne by large users. Let $1-\psi \in (0,1)$ denote the asymmetry in tax burden, i.e., if $\psi = 0$ the burden falls exclusively on small users, whereas it is shared evenly if $\psi = 1$. Then the objective of the central government as a function of the grant allocation $G_1, G_2, \ldots$ to different communities

reduces to

$$V^c(G_1, G_2, \ldots) \equiv \quad \Sigma_i N_i \mathcal{E}_{\eta_i,\theta_i}[\beta_i\gamma_s\eta_i v(f_s Y_i^L) + \delta^c\{(1-\beta_i)\gamma_l\eta_i v(f_l Y_i^L) + G_i - \theta_i Y_i^L\}$$
$$-\{1+\psi(\delta^c-1)(1-\beta_i)\}(1+\lambda)G_i],$$

which can be expressed as the sum of separate objective functions for different regions:

$$V^c(G_1, G_2, \ldots) \equiv \Sigma_i N_i V_i^c(G_i) \tag{22}$$

where the objective function corresponding to community i is a function of the grant to that community alone:

$$V_i^c(G_i) \equiv \quad \mathcal{E}_{\eta_i,\theta_i}[\beta_i\gamma_s\eta_i v(f_s Y_i^L) + \delta^c\{(1-\beta_i)\gamma_l\eta_i v(f_l Y_i^L)$$
$$+G_i - \theta_i Y_i^L\} - \{1+\psi(\delta^c-1)(1-\beta_i)\}(1+\lambda)G_i] \tag{23}$$

Therefore the community grant G_i will be selected to maximize (23). The central government therefore incorporates its awareness of how how any given grant to a community will tend to be allocated (or diverted) within that community, in deciding on the level of the grant.

The analysis of optimal community grants is somewhat complicated, and so we omit some of the technical steps (which are available in the working paper version of this paper Bardhan and Mookherjee (2000c)). The overall implications are summarized below.

Proposition 5 *With decentralization financed entirely by central grants:*

(a) Region i will be financially constrained with positive probability if

$$(1+\lambda) > [\frac{1}{\delta^c} + \psi(1-\beta_i)(1-\frac{1}{\delta^c})]^{-1} \tag{24}$$

In this case, region i will be financially constrained if and only if local need shock η_i exceeds the threshold η_i^*.

(b) In low need states where region i is not financially constrained, large users are provided first-best service levels (besides the benefits of diverted funds), while small users are underprovided relative to the first-best to an extent depending on local capture. In financially constrained states, service levels are the same as at the threshold state η_i^*, and no funds are diverted.

(c) Service delivery levels for either group is smaller in all states compared with user-fee-financed decentralization.

(d) If the deadweight cost of taxes λ is sufficiently large, grant-financed decentralization is less efficient than centralization as well as decentralization financed by central taxes or user fees. If λ is sufficiently small, local and central capture $(\delta_i^l - 1), \psi(\delta^c - 1)$ sufficiently close to zero, then grant-financed decentralization approaches the first-best.

Note the importance of financing constraints faced by the central government, represented by λ. Even with perfectly accountable governments, financing constraints at the central level will lead to service underprovision with grant financing, unlike decentralization based on local financing or user fees. As λ rises, service levels will progressively shrink as central grants dry up. For λ sufficiently large, service levels will decline precipitously, causing performance to drop below centralization as well. At the other extreme, if collection at the center is efficient and λ is close to zero, and governments are sufficiently accountable at both levels, grant-financed decentralization will tend to first-best efficiency.

In particular, note that grant-financing may be dominated by user-fee financing under appropriate conditions (e.g., λ sufficiently large), while under others grant financing may be more or almost as efficient than user fees. To gain further insight into the relevant trade-offs, we compare the resulting patterns of service deliveries with centralization and user-fee-financed decentralization. These are

depicted in Figures (4) through (7). Figure 4 compares deliveries with those under centralization for regions where $\beta_i < \beta_1^*$, where the threshold β_1^* is defined by the condition $\delta_i^l = [1 - \frac{1-\beta_1^*}{\beta_i^*}(\frac{\gamma_l}{\gamma_s} - 1)]^{-1}$, i.e., service undeprovision to small users is the same as in centralization. In low β_i regions, grant financing expands supplies to small users in low need states where the local government is not financially constrained, while supplies to large users is unaffected. But in high-need states where local financing constraints bind, service levels may shrink for both groups under grant financing. The overall effect on service levels and efficiency thus depends on the severity of the local financing constraints, as explained above. One apparent benefit of decentralization in these regions is that it improves equity in service levels (in low need states); this may however not be mirrored in a genuine improvement in equity since small users may ultimately bear a greater financial burden under decentralization (e.g., if they bear a disproprtionate share of the burden of financing central taxes).

For regions where the fraction of small users β_i is larger (depicted in Figure 5), on the other hand, service allocations to small users shrink under decentralization even when local governments are financially constrained, and equity in service allocation deteriorates. In this case, service levels shrink for both classes of users (except large users in low need states, who are served the same). Here grant-financed decentralization hurts growth, efficiency as well as equity. Indeed, it may be Pareto-inferior to centralization: large users may be worse off even if they bear a negligible fraction of the burden of central taxes, as a consequence of shrinking central grants that dry up service deliveries. This is case where residents will regret the absence of the corruption that 'lubricated' the centralized system in the past!

We turn finally to the comparison with user-fee financed decentralization, which we have shown above dominated local-tax-financing. Note that user-fee financing generates efficient provision of

the service if and only if β_i is small enough, whence the burden of financing over-provision on small users tends to be excessive. For instance in the case where α, the elasticity of marginal utility of consumption, lies between 0 and -1, the relevant threshold β_2^* is defined by the condition $\frac{\beta_2^*}{1-\beta_2^*}[\frac{1}{1+\alpha} - 1] = (\frac{\gamma_s}{\gamma_l})^{\frac{1}{\alpha}}$. For β^i smaller than this threshold, both categories of users are served at the first-best level under user-financing, and small users bear the entire fiscal burden. For β^i higher, large users are overprovided, while small users continue to be efficiently funded. See Figures 6 and 7 for these two cases respectively. It therefore follows that in all cases, grant financing shrinks service levels to both categories of users relative to user-fee financing, with the exception of large users in low β_i regions when their local governments are not financially constrained. Hence *the effect of not devolving revenue raising powers to local goverments in step with their expenditure responsibilities causes an unambiguous reduction in the level of services in all regions, irrespective of patterns of political accountability.* This is again consistent with the empirical results of Estache and Sinha (1995). The reduction in service levels may however be efficiency enhancing, as they constrain the tendency for large users to be overprovided under user-fee financing. Since such overprovision is paid for by the small users, it may improve local equity as well. The problem with grant-financing on the other hand is the tendency for small users to be under-served if local governments are susceptible to capture, or both categories to get under-served in high-need states when financing constraints bind. The severity of the latter problems depend respectively on the extent of local capture, and on the costs of raising central taxes. If they are not very acute, grant-financing may conceivably end up dominating user-fee financing.

5 Concluding Comments

This paper has studied the tradeoffs between allocation distortions resulting from monopoly power of unregulated and corrupt bureaucrats in a centralized delivery system, and the tendency for local governments to be captured by local elites under decentralization. The key point is that the effects of decentralizing service delivery will depend on the method chosen for financing local governments. Existing empirical results suggest that expenditure decentralization not accompanied by revenue decentralization limit the expansionary effect of decentralization on service levels. Our model provides an explanation for this pattern, and at the same time urges caution in inferring that greater revenue decentralization would be welfare enhancing. Local capture tends to be manifested in service overprovision to local elites, at the expense of elites, which is both inefficient and inequitable. Accordingly restraints on the revenue-raising capability of local governments can limit the extent of such resource misallocations.

User fee financing mechanisms are particularly notable in this connection: the voluntariness of such mechanisms in contrast to the coercive character of local taxes limit the extent of regressive redistributions that elites can employ in their favor. Indeed, in our model user fees ensure that decentralization is welfare enhancing relative to centralization, irrespective of the extent of local capture. Compared with the more traditional form of financing, i.e., intergovernmental fiscal grants, user fees have the added advantage of fiscal self-sufficiency of local governments. This enables service allocations to be sensitive to random fluctuations in local costs and needs, particularly when such flexibility is most useful (i.e., when local need is high). They also ensure higher service deliveries compared with grant financing, owing to the avoidance of asymmetric information, inter-community free-riding and bargaining distortions inherent in a system of intergovernmental fiscal grants.

Apart from the normative results, our model also provides a number of detailed predictions

concerning the impact of decentralization on service allocations and their financing, which are empirically testable. We hope that future empirical analyses of fiscal decentralization in developing countries will be carried out to test these predictions.

Our model abstracted from problems of interregional spillovers of decisions made by local governments, and possible lack of expertise at local levels. Both of these may be important in practice, and need to be evaluated independently in assessing the effects of decentralization. Spillovers might naturally arise in the areas of roads, telecommunications, schools and public health. Even in the context of water resources, spillovers will arise if there are capacity constraints on the amount of the service that can be produced and delivered across all regions. In all of these cases, decentralization will require coordination of decisions made independently by different local governments, involving either central interventions, establishment of resource sharing formulae, or market-like mechanisms. Lack of managerial and technical expertise at the local level may prevent cost-effective provision of the service within regions. For instance, if local government officials are not informed about the realization of marginal costs of serving their community, managers of the central production enterprise may be able to earn rents by exploiting their specialized information, resulting in additional distortions under decentralization. Cost-effective procurement may also be vitiated if local government officials do not have much bargaining power when dealing with service providers, allowing the latter to earn monopoly rents. In the presence of either of these problems, the performance of decentralized regimes will deteriorate further; the results of this paper accordingly ought to be qualified.

We considered three polar modes of financing most commonly employed in developing countries (see, e.g., (Dillinger (1995)). Mixed modes of financing may also be worth exploring in this context, e.g., where local governments rely on a mixture of local user charges and central grants, which might dominate either polar mode. We also restricted attention to unrestricted fiscal grants, which

allowed some degree of flexibility in local service delivery, at the cost of allowing diversions of surplus resources to local elites in low-need states. Grants tied to expenditures on specific services restrict both flexibility in service delivery and scope for diversion of unspent funds to other less important social purposes. An intermediate form of grant finance involves matching grants tied to specific services, which combine advantages of providing some degree of flexibility in service delivery, while limiting scope for diversion. Clearly the welfare implications of a richer set of financing options than analysed in this paper deserve to be explored in future research.

References

Ahmad E. and N. Stern (1984), "The Theory of Tax Reform and Indian Indirect Taxes", *Journal of Public Economics,* 25(3), 259-298.

H. Alderman, "Do Local Officials Know Something We Don't? Decentralization of Targeted Transfers in Albania", Washington DC, World Bank, 1999.

A. Banerjee, "A Theory of Misgovernance," *Quarterly Journal of Economics*, 62, 1289- 1332, 1997.

P. Bardhan, "Efficiency, Equity and Poverty Alleviation: Policy Issues in Less Developed Countries", *Economic Journal*, vol. 106, September 1996a.

——— "Decentralized Development", *Indian Economic Review*, vol.31, no.2, 1996b.

——— and Dilip Mookherjee, "Relative Capture of Local and Central Governments: An Essay in the Political Economy of Decentralization," Working Paper, Institute for Economic Development, Boston University, 1999.

————————-, "Capture and Governance at Local and National Levels," *American Economic Review*, forthcoming, 2000a.

———————–, "Decentralizing Anti-Poverty Program Delivery in Developing Countries," mimeo, Institute for Economic Development, Boston University, 2000b.

———————, "Corruption and Decentralization of Infrastructure Delivery in Developing Countries," mimeo, Institute for Economic Development, Boston University, 2000c.

D. Baron and R. Myerson, "Regulating a Monopolist with Unknown Cost," *Econometrica*, 1982.

D. Baron, "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review*, 88, 1994, 33-47.

T. Besley, "Targeting Taxes and Transfers: Administrative Costs and Policy Design in Developing Economies," Development Studies Working Paper 146, 1989, Princeton University, Woodrow Wilson School.

——— and R. Kanbur, "Principles of Targeting," in Michael Lipton and Jacques ven der Gaag, eds., *Including the Poor*, Washington DC, World Bank, 1993.

——— and S.Coate, "Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Analysis," National Bureau of Economic Research Working Paper no. 7084, April 1999.

R. Bird, "Decentralizing Infrastructure: For Good or Ill?" in A. Estache (ed.), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Papers 290, Washington D.C., 1995.

P. Bolton and G. Roland, "The Breakup of Nations: A Political Economy Analysis," *Quarterly Journal of Economics*, 62, 1057-1090, 1997.

J. Conning amd M. Kevane, "Community Based Targeting Mechanisms for Social Safety Nets," mimeo, Williams College, 1999.

A. Coudouel, S. Marnie. et. al., Targeting Social Assistance in a Transition Economy: The Mahallas in Uzbekistan, UNICEF, 1998.

J. Cremer, A. Estache, and P. Seabright, "The Decentralization of Public Services: Lessons from the Theory of the Firm", in A. Estache (ed.), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Papers 290, Washington D.C., 1995.

medskip

Richard Crook and James Manor, *Democracy and Decentralisation in South Asia and West Africa*, Cambridge: Cambridge University Press, 1998.

A. Das-Gupta and D. Mookherjee, *Incentive and Institutional Reforms in Tax Enforcement: An Analysis of Developing Country Experience*, New Delhi, Oxford University Press, 1998.

B. Dillinger, "Decentralization, Politics and Public Service," in A. Estache (ed.), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Papers 290, Washington D.C., 1995.

J. Dreze and M. Saran, "Primary Education and Economic Development in China and India: Overview and Two Case Studies", in K. Basu et al (eds.), *Choice, Welfare, and Development: A Festchrift in Honor of Amartya K. Sen*, Clarendon Press, Oxford 1995.

A. Estache (Editor), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Paper no. 290, 1995.

——— and S. Sinha, "Does Decentralization Increase Public Infrastructure Expenditure?" in A. Estache (ed.), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Papers 290, Washington D.C., 1995.

R. Fisman and R. Gatti, "Decentralization and Corruption: Evidence Across Countries," Working Paper,

Development Research Group, World Bank, October 1999.

E. Galasso and M. Ravallion, "Distributional Outcomes of a Decentralized Welfare Program," 1999, mimeo, World Bank, Washington DC.

M.E. Grosh, "The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions," LSMS Working Paper 80, the World Bank, 1991.

——, "Towards Quantifying the Tradeoff: Administrative Costs and Targeting Accuracy," in van de Walle and Nead (ed.), 1995.

G. Grossman and E. Helpman, "Electoral Competition and Special Interest Politics," *Review of Economic Studies*, 63, 265-286, 1996.

F. Humplick and A. Estache, "Does Decentralization Improve Infrastructure Performance?" in A. Estache (ed.), *Decentralizing Infrastructure: Advantages and Limitations*, World Bank Discussion Papers 290, Washington D.C., 1995.

R. P. Inman and D. L. Rubinfeld, "Designing Tax Policies in Federalist Economies: An Overview", *Journal of Public Economics*, vol.60, 1996.

————————————-, "Rethinking Federalism," *Journal of Economic Perspectives*, 11(4), 1997, 43-64.

T.M. Thomas Isaac and Richard Franke, *Local Democracy and Development: People's Campaign for Decentralised Planning in Kerala*, Left Word Books, New Delhi, 2000.

E. Jimenez and Y. Sawada, "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program," *World Bank Research Observer*, 1999, 13(3), 415-441.

J.J. Laffont and J. Pouyet, "The Subsidiarity Bias in Regulation," mimeo, University of Toulouse, 2000.

G K Lieten, "Panchayats in Western Uttar Pradesh," *Economic and Political Weekly*, September 28 1996, 2700-2705.

M. Lipton and M. Ravallion, "Poverty and Policy," in *Handbook of Development Economics*, Volume III, edited by J. Behrman and T.N. Srinivasan, Chapter 41, Amsterdam: North-Holland, 1995.

B. Lockwood, "Distributive Politics and the Benefits of Decentralization," 1998, CSGR Working Paper No. 10/98

George Mathew and Ramesh Nayak, "Panchayats at Work: What It Means for the Oppressed?" *Economic and Political Weekly*, July 6 1996, 1765-1771.

R. Musgrave and P. Musgrave, *Public Finance in Theory and Practice*, 4th Edition, New York, McGraw Hill, 1984.

D. Newbery and N. Stern, *The Theory of Taxation for Developing Countries,* 1987, World Bank: Oxford University Press.

Boaventura de Sousa Santos, "Participatory Budgeting in Porto Alegre: Toward a Redistributive Democracy," mimeo, Centro de Estudios Sociais, Coimbra, Portugal, 1997.

W. E. Oates, *Fiscal Federalism*, Harcourt, Brace and Jovanovich, New York, 1972.

P. Seabright, "Accountability and Decentralization in Government: An Incomplete Contracts Model", *European Economic Review*, vol. 40, 1996.

M. Tommasi and F. Weinschelbaum, "A Principal-Agent Building Block for the Study of Decentralization and Integration," mimeo, Universidad de San Andres, 1999.

D. van de Walle and K. Nead (eds.), *Public Spending and the Poor: Theory and Evidence*, Johns Hopkins University Press, Baltimore, 1995.

Robert Wade, "How Infrastructure Agencies Motivate Staff: Canal Irrigation in India and the Republic of Korea," in *Infrastructure Strategies in East Asia*, Ashoka Mody (ed.), World Bank, 1997.

L. Wong , "Privatization of Social Welfare in Post-Mao China," *Asian Survey*, 1994, 34(4): 307-325.

World Bank, *Entering the 21st Century*, World Development Report, 2000, forthcoming, The World Bank, Oxford University Press.

Figure 1: SERVICE DELIVERY PATTERNS UNDER CENTRALIZATION AND LOCAL-TAX-FINANCED DECENTRALIZATION: $1 + \lambda = \frac{\eta_i}{J(\eta_i)}$

Figure 2: SERVICE DELIVERY PATTERNS UNDER USER-FEE-FINANCED AND LOCAL-TAX-FINANCED DECENTRALIZATION

Figure 3: PER CAPITA SERVICE DELIVERY WITH GRANT FINANCING

Figure 4: SERVICE PATTERNS UNDER CENTRALIZATION AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i < \beta_1^*$

Figure 5: SERVICE PATTERNS UNDER CENTRALIZATION AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i > \beta_1^*$

Figure 6: SERVICE PATTERNS UNDER USER-FEE-FINANCED AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i < \beta_2^*$

Figure 7: SERVICE PATTERNS UNDER USER-FEE-FINANCED AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i > \beta_2^*$

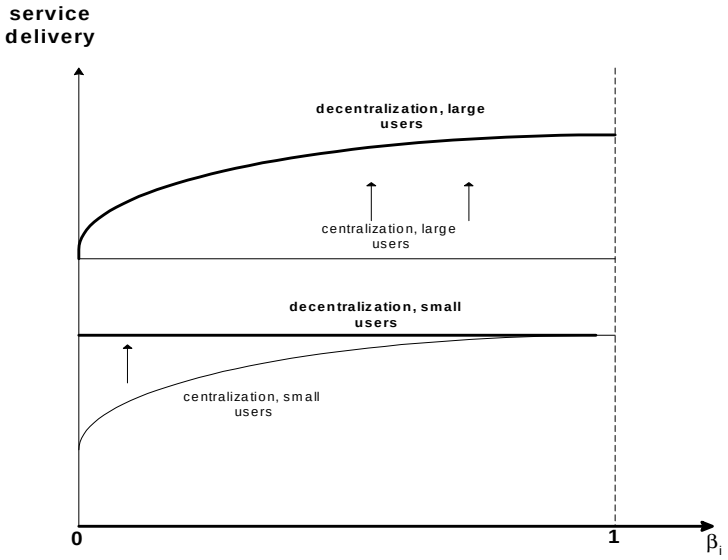

Figure 1: SERVICE DELIVERY PATTERNS UNDER CENTRALIZATION AND LOCAL-TAX-FINANCED DECENTRALIZATION: $1 + \lambda = \frac{\eta_i}{J(\eta_i)}$

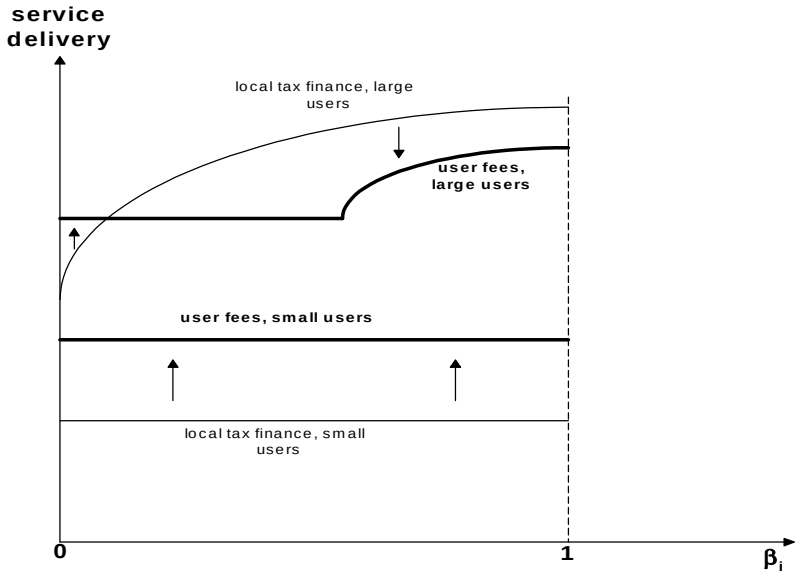

Figure 2: SERVICE DELIVERY PATTERNS UNDER USER-FEE-FINANCED AND LOCAL-TAX-FINANCED DECENTRALIZATION

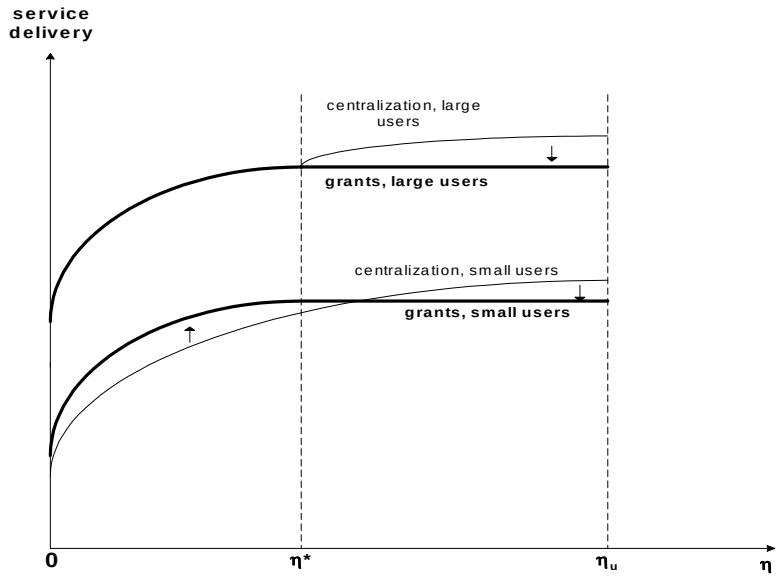

Figure 3: SERVICE PATTERNS UNDER CENTRALIZATION AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i < \beta_1^*$

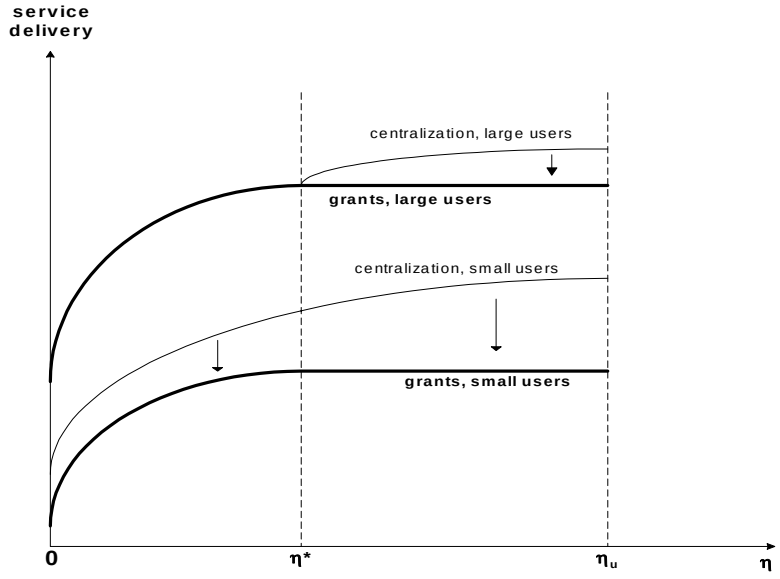

Figure 4: SERVICE PATTERNS UNDER CENTRALIZATION AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i > \beta_1^*$

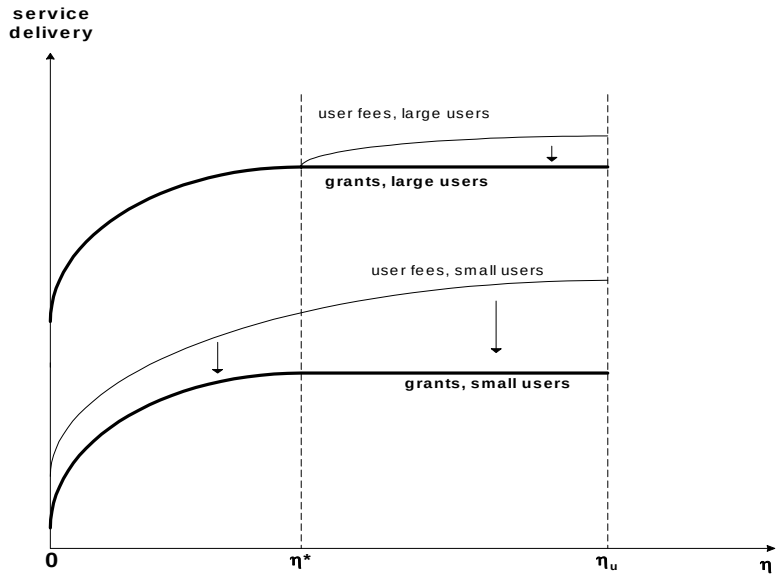

Figure 5: SERVICE PATTERNS UNDER USER-FEE-FINANCED AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i < \beta_2^*$

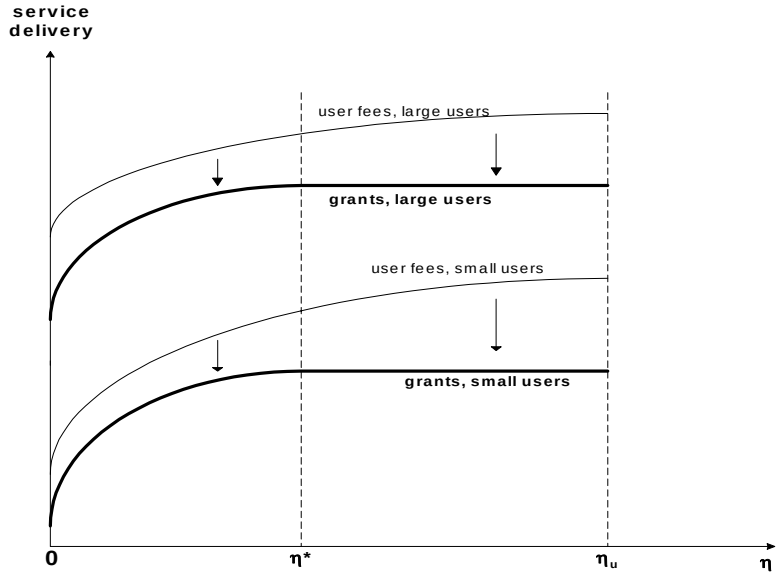

Figure 6: SERVICE PATTERNS UNDER USER-FEE-FINANCED AND GRANT-FINANCED DECENTRALIZATION IN REGIONS WITH $\beta_i > \beta_2^*$