

Decentralizing Anti-Poverty Program Delivery in Developing Countries

Pranab Bardhan

Department of Economics, University of California, Berkeley

and

Dilip Mookherjee

Department of Economics, Boston University

Abstract

We study effects on accountability in government service delivery of decentralizing administration of an antipoverty program. While governments at both central and local levels are vulnerable to anti-poor policy biases owing to political capture, centralized delivery systems are additionally prone to bureaucratic corruption owing to problems in monitoring bureaucratic performance. Decentralizing the delivery system promotes cost-effectiveness and improves intra-regional targeting at low program scales. But inter-regional targeting may deteriorate, as central grants to high poverty regions shrink, owing to high capture of local governments by local elites in such regions.

Keywords: bureaucracy, corruption, decentralization, poverty alleviation, targeting

JEL Classification Nos.: D72, D73, H 41, H42, H77, I38, O17

Corresponding Author: Dilip Mookherjee, Department of Economics, 270 Bay State Road, Boston MA 02215; tel no: 617-3534392, fax no: (617)3534392, dilipm@bu.edu

1 Introduction

Decentralization of authority over administration of redistribution programs to local communities has recently become widespread in the developing world. These initiatives have transferred responsibility from central ministries to local governments or community representatives over procurement, selection of local projects and identification of beneficiaries. Such experiments have been introduced since the 1980s in Armenia, Albania, Bosnia-Herzegovina, Brazil, China, El Salvador, Georgia, India, Mexico, South Africa, Uganda and Uzbekistan (Conning and Kevane (1999), Jimenez and Sawada (1999), Mookherjee (2001) and World Bank (2000)). Other countries still characterized by traditional top-down delivery systems are currently contemplating adopting similar reforms.

The principal aim of these reforms has been to enhance accountability in the service delivery process by limiting problems inherent in the traditional centralized system. One of the most important of these is corruption among centrally appointed bureaucrats allocated authority over service delivery in local areas. While systematic empirical evidence on corruption is scarce for obvious reasons, there are innumerable case studies of high rates of corruption among government bureaucrats in the process of delivering public services.¹

The presumed argument in favor of decentralizing the delivery system is that local governments will be subject to electoral pressures from local citizens who are able to monitor delivery better than a distant central authority. Large scale bribery, cost-padding and diversions to the black market will be more noticeable by local residents that have to pay

¹Wade (1985, p. 474)) for instance describes the bribery among irrigation bureaucrats in the Indian state of Andhra Pradesh, where he estimates illicit earnings from water bribes to be one or two times the annual salary for an Assistant Engineer, and three to five times for an Executive Engineer. Rose-Ackerman (1999, Chapter 1) cites numerous reports concerning diversion of commodities to the black market, and charging of bribes for delivering public services by government bureaucrats in diverse countries. These include diversion of coal and other producer goods to the black market in China at a profit of between 250 to 680%; payoffs needed to obtain subsidized loans in Eastern Europe, Russia, Lebanon, Kenya, Fiji, Korea and Pakistan; bribes for certifying unqualified people for subsidized housing in Washington D.C.; corrupt payments for receiving the best plots in land reform programs in India and Pakistan; fraud and corruption in processing of workmen's compensation claims in Brazil; 'routine' payments that need to be made by citizens to receive telephone installations, driving licenses, passports, electric connections in India, Russia and Ukraine.

these bribes and live in close proximity to local government officials. Residents are then likely to be motivated to eject such government officials from office, besides applying social sanctions. This will conceivably enhance accountability in the delivery process.

Yet this presumption of greater accountability under decentralization is frequently questioned, since it is based on the assumption that local democracy will function effectively. It assumes a level of literacy and political awareness among citizens that may be unrealistic for many poor countries. Numerous case studies exist of development programs being stymied by capture of local governments by powerful local elites, that distort and divert public programs to benefit themselves at the expense of poor minorities.² In the *Federalist Papers* (no. 10), such concerns about local capture led James Madison to argue for retention of powers at the federal level (Hamilton, Madison and Jay (1787[1937 pp 62–70])). Political scientists (such as Crook and Manor (1998) have accordingly studied patterns of participation by the poor in local politics in a number of countries in South Asia and Africa, in order to predict the level of accountability that can be expected from local governments under decentralization. These empirical studies, as well as our own theoretical analyses (Bardhan and Mookherjee (1999, 2000)) lead to no uniform or general conclusions about relative vulnerability of different levels of government to special interest capture.³

The purpose of this paper is to treat the extent of local capture (*vis-a-vis* capture of

²Gaiha, Kaushik and Kulkarni (1998, 2000) conclude following a survey of implementation of rural development programs implemented by local government bodies (called *gram panchayats*) in the state of Uttar Pradesh (UP), India:

”A survey of 3 districts in UP on the implementation of the 2 major anti-poverty programs, viz. JRY and IRDP in 1997 points to the pervasiveness of rent-seeking behavior in all the Gram Panchayats. As a result, there were large-scale diversions of resources. The benefits of JRY, IRDP and the other rural development programs to the poor were minimum.... Regardless of whether the Panchayat chairperson belonged to upper or lower castes, there was little variation in the method and pattern of domination by the locally powerful groups”.

In similar vein, Tandler (1997) describes how powerful local elites tended to be the most important roadblock in improving health services in the Brazilian town of Ceara.

³This is in contrast to the approach of Seabright (1996), who assumes that accountability is invariably higher at the local level. In his theory this is a benefit of decentralization which needs to be traded off against possible losses in policy coordination across regions.

higher level governments) as an exogenous parameter reflecting local poverty and political institutions. Our purpose is to theoretically examine their consequences for the impact of decentralization on potentially observable measures of performance of an anti-poverty program. These measures include cost-effectiveness, community insurance against shocks to local costs or need, intra-regional and inter-regional targeting.

We model the implications of a corrupt centralized bureaucracy that is either not subject to any monitoring at all by their supervisors in the central ministry, or to weak auditing and oversight mechanisms. The lack of control over bureaucrats' performance in the centralized system stems from problems of measuring their performance, owing to the absence of suitable market-based metrics, and high communication costs endemic in developing countries. Efforts to monitor targeting performance of merit goods in far-flung communities in a country with poor accounting and information systems are quixotic at best. Consequently, centrally appointed bureaucrats with monopoly power over local service delivery face few obstacles in padding costs and diverting goods to the black market. For the sake of theoretical clarity we assume for most part that central bureaucrats are subject to no oversight at all. The principal results are qualitatively unaltered if bureaucrats are subject instead to random infrequent leaks or audits that result in imposition of (pecuniary or nonpecuniary) penalties.⁴

In the decentralized system, delivery authority is shifted to elected local governments. These governments are subject to electoral competition with special interest capture, in the style of Grossman and Helpman (1996). The local government is responsive to the welfares of local residents, but with a bias against the poor. This represents the phenomenon of local capture, which arises for two reasons: lower levels of political awareness of the nonpoor, and campaign contributions by the nonpoor in exchange for preferential treatment in services. Political influence replaces bribery as the mechanism for diverting services to local elites

⁴The problem with centralization that we focus on, therefore, has relatively little to do with corruption at higher levels of government. Central politicians may or may not be able to receive some kickbacks from central bureaucrats: the key point is that they hardly have any power to control the corruption even if they wished to. Accordingly informed voters cannot expect central politicians to be able to root out corruption: the targeting failures are endemic to a centralized system, irrespective of whichever party happens to be in power.

under decentralization. The extent of capture of a local government is likely to be higher in communities with a higher poverty rate.

Higher levels of government are also subject to a similar process of electoral competition on a national scale. Consequently both local and national governments will have an anti-poor bias, though possibly to varying degrees (which we treat as a parameter). The national government is responsible for raising revenues and allocating lumpsum grants to different communities, based only on information concerning population sizes and poverty rates of different regions (but not the realization of local needs or costs). Inter-regional allocations will thus depend on the political objectives of the national government, and on its expectation regarding the nature of intra-regional targeting of allocated grants under either centralized or decentralized systems.

In this setting it is not surprising that the overall welfare effect of decentralization will depend on the extent of capture of local governments. Of greater interest are the implications for different observable dimensions of performance. In this respect we derive a number of results which are qualitatively robust with respect to different specifications of the model:

- (i) The decentralized regime is in general more cost-effective, owing principally to the fact that local governments represents the interests of users, in contrast to bureaucrats appointed by the central government.
- (ii) Neither regime is able to provide for any degree of community insurance, owing to incentive problems between local and central government officials.⁵
- (iii) With respect to intra-regional targeting, the decentralized regime performs better at smaller scales of the program, but worse at higher scales. Owing to this, the center constrains grants to local governments in the decentralized system to scales

⁵This result stems from our assumption that local governments do not have any authority or ability to raise local revenues. In the presence of local revenue raising ability, the decentralized regime is capable of providing some degree of community insurance under a system of matching grants from the central government. It also depends on our abstraction from self-targeting mechanisms, where distributions of the good are bundled with a work requirement.

small enough to prevent leakages. The consequence is that the *realized* intra-regional targeting performance of the decentralized system is superior.

- (iv) The principal drawback with decentralization in regions of poor local governance will thus be manifested in a smaller scale of the program (implying greater cutbacks to the poor). Hence intra-regional targeting is likely to deteriorate: wealthier regions will witness greater reductions in poverty, creating greater inter-regional disparities in poverty.

It is important to emphasize the distinction between our approach and traditional theories of fiscal federalism (e.g., Oates (1972), and reviewed recently in Inman and Rubinfeld (1996, 1997), and Besley and Coate (1999)). Our principal focus is on effects of decentralization on accountability in the service delivery process, an issue typically ignored in traditional approaches. In line with the reality of most poor countries, we abstract away from the possibility of citizens ‘voting with the feet’ as a way of enforcing some accountability in the centralized system. For the sake of theoretical clarity, we use a framework which is the opposite extreme of a Tiebout-style model and assume that every community has a heterogenous composition, with mobility costs high enough to prohibit any movements of people or goods across communities.⁶ We also abstract from inter-community externalities, which also seems reasonable enough for many antipoverty programs involving distribution of food, clothing or housing generally, besides education or health programs when there is little movement across communities.

Our results can also be contrasted to those of the traditional approach. For example, almost all versions of the traditional literature agree on the principle that local public goods not subject to economies of scale or inter-regional spillovers should invariably be provided by local governments. Heterogeneity across communities in our approach weakens the ar-

⁶There is only limited evidence of Tiebout-style ‘voting with the feet’ or ‘welfare magnet’ effects even in the case of developed countries, as reported by numerous studies cited in Conning and Kevane (1999, p. 32). This assumption is even more realistic in developing countries, owing to higher mobility and transport costs. These include a greater importance of informal social norms that differentiate sharply between ‘outsiders’ and ‘insiders’, making it more difficult for a resident of one region to be entitled to welfare benefits in other regions.

gument for decentralization (owing to the poorer inter-regional targeting performance that results), in contrast to the traditional approach where it strengthens the case for decentralization (owing to the greater diversity it allows relative to centralization). Moreover, the form of decentralization we study is far more limited than is typically considered in the traditional literature, which includes devolution of taxing authority to local governments. It is more accurate to view the recent trends witnessed in developing countries as delegation of service delivery by the central government to a different set of agents, rather than a shrinkage in its authority. Hence many of the policy concerns in the traditional literature on fiscal federalism are not pertinent, e.g., concerning the need to coordinate tax policies across regions. Instead our principal focus is on cost-effectiveness and targeting of service delivery within regions. The nature of the trade-offs involved, and the policy implications, are therefore quite distinct.

It is worthwhile to clarify the essence of the contrast between centralization and decentralization in our model. The nature of the incentive problem is qualitatively different across the two systems: authority is shifted away from unsupervised central bureaucrats that seek to maximize bribe incomes, towards elected officials of local governments whose objective is to maximize the chances of staying in power. Economic corruption tends to be replaced by political corruption (in the form of diversion of services to local elites), despite the fact that all agents in either system behave in a self-interested fashion. The difference between the two systems does not stem from an *ad hoc* disparity between underlying preferences of concerned decision-makers, but rather from differing patterns of accountability arising from different information and incentive systems.⁷

The nature of endowments, preferences and technology is presented in Section 2. Section 3 describes the nature of political capture that either local or central governments are subject

⁷By contrasting communication costs that reduce the capacity of centralized systems to devise effective incentive systems, with the incentive problems inherent in delegating decision-making to local agents, our theory is similar to models of organizations developed by Green and Laffont (1986, 1987), Melumad, Mookherjee and Reichelstein (1992, 1997), and Laffont and Martimort (1998). Banerjee (1997) presents a related theory of targeting failures in a centralized system arising from bureaucratic corruption and wealth constrained citizens. Mookherjee (1998, 2001) elaborates on similar problems in the design of incentive systems for bureaucrats in developing countries.

to, and Section 4 the problem of bureaucratic corruption that the centralized system is prone to. Sections 5 and 6 respectively describe the outcomes of the centralized and decentralized systems. Section 7 examines their relative performance with respect to targeting, in a setting which abstracts from cost-effectiveness and community insurance. Finally, Section 8 concludes.

2 Endowments and Utilities

There are m communities $i = 1, 2, \dots$. Community i has N_i citizens, and is comprised of two classes: the poor and nonpoor.⁸ Let the proportion of poor in community i be denoted π_i . There are two goods: a consumption good, and money. A member of class $k (= p, n)$ in community i has a utility function $v_k(y; \eta_i) + x$ defined over external supply (or purchase) y of the good by that person, and x , money left over after paying for purchases of this good. The function v_k is smooth, strictly increasing and concave in y . Moreover, the good — like milk, tap water, vitamin supplements or schooling — is a necessity but not essential to survival. Hence total and marginal utility at zero consumption are both finite, while market demand for the nonpoor is subject to zero income elasticity, and a price elasticity less than one.

A key but natural assumption is that at any given level of service delivery, the poor derive a higher marginal utility, since they have smaller endowments of these goods to start with. Moreover, both total and marginal utility is strictly increasing in η_i , a parameter of local need that applies to all members of community i (except at the point of zero consumption where η_i does not affect the utility level). The realization of η_i is known to all members of the community, but unknown to the central government, whose beliefs are represented by a probability distribution over a range $[\underline{\eta}_i, \bar{\eta}_i]$.

The good is available on a perfectly competitive market within each community. Suppliers of the good in community i incur a unit cost θ_i of delivering this good, independent of

⁸The nonpoor could be heterogenous with respect to their assets and preferences: as we shall see this will not matter, as long as their market purchases of the relevant good are not characterized by (local) income effects. We will therefore simplify by assuming that members of each class are homogenous.

the quantity delivered, where θ_i has a given probability distribution over the range $[\underline{\theta}_i, \bar{\theta}_i]$. Owing to competition among suppliers, this good is supplied in community i at a constant price of θ_i . The local bureaucrat and local residents can costlessly resell the good to others within their own community, so local black markets can open. In order to avoid complications arising from market links across communities, we assume that they cannot participate in the black markets of other communities (e.g., because the good in question is location-specific, or owing to large transport costs).⁹

The poor have no money at all (after meeting their subsistence requirements), and can survive without purchasing any of the good, corresponding to the assumption that $v_p(0, \eta_i)$ is finite. So they will be not to able to purchase the good on the market, and end up with a utility of $v_p(0, \eta_i)$ in the absence of any government intervention. The nonpoor in contrast have sufficient money that their market purchases of the good are independent of their money stocks. Each nonpoor person consumes $y_n^*(\theta_i, \eta_i)$ units of the good, where $v'_n(y_n^*(\theta_i, \eta_i), \eta_i) = \theta_i$, achieving indirect utility $v_n(y_n^*(\theta_i, \eta_i), \eta_i) - \theta_i y_n^*(\theta_i, \eta_i)$ plus their initial money stocks. Since the good is a necessity, we shall assume that

$$v'_n(0, \underline{\eta}_i) > \bar{\theta}_i, \tag{1}$$

so every nonpoor person always buys positive quantities, but purchases are price inelastic (the elasticity of y_n^* with respect to θ_i is less than one).

The constrained wealth of the poor prevents competitive markets from maximizing social surplus: the marginal utility of the good is higher for the poor, yet they are unable to enter into trades with the nonpoor who have a lower marginal utility for the good. This is precisely the welfare rationale for public provision of the good to the poor, financed by taxes levied on the nonpoor. If the deadweight cost per dollar taxes imposed on the non-poor is λ , then

⁹Vouchers distributed by the program can be location-specific, as in the case of housing, schooling or health services. In the case of food, however, the good can conceivably be sold outside the local community. But even this may entail significant transaction and transport costs, which may not make it worthwhile for individual households in isolated communities. Of course, if the financial returns are large enough, intermediaries could form to economize on these costs. In cases where sales in external markets is feasible, it will turn out that the performance of the centralized system will become worse. A more detailed examination of this case may be worthwhile in future work.

from an ideal utilitarian standpoint setting up such a redistributive program in community i would be worthwhile as long as (using \mathbf{E} to denote the expectation operator):

$$1 + \lambda < \frac{\mathbf{E}_{\eta_i} v_p'(0, \eta_i)}{\mathbf{E} \theta_i}, \quad (2)$$

assuming that the program purchases the good at cost θ_i and distributes it to the poor in community i . However, this presumes a number of ideal conditions: that the good is purchased cost-effectively and targeted perfectly, and the government in question is occupied by a benevolent planner seeking to maximize a utilitarian welfare function which assigns identical welfare weights to the poor and the non-poor.¹⁰

To illustrate the problems raised by lack of cost-effectiveness, community insurance and imperfect targeting, consider a hypothetical scenario where there is neither any political capture nor bureaucratic corruption. Then the government maximizes the utilitarian welfare function:

$$\sum_{i=1}^m N_i [\pi_i v_p(y_{pi}, \eta_i) + (1 - \pi_i) \{v_n(y_{ni}, \eta_i) - t_i\}] \quad (3)$$

where y_{ki} denotes the consumption of a class k person in community i , and t_i is the net financial cost incurred by the nonpoor as a result of the program. The government would have access to all relevant information and face no implementation problems, and the ideal *first-best allocation* would result: poor individual in a community with local need η_i and cost θ_i would be delivered a quantity $y_1(\theta_i, \eta_i)$ of the good that would equate his marginal utility of consumption to the true cost of the good, inclusive of the deadweight costs of the tax revenue required to finance its provision:

$$\frac{1}{\theta_i} v_p'(y_1(\theta_i, \eta_i), \eta_i) = (1 + \lambda). \quad (4)$$

Let the corresponding aggregate supply and expenditure grant for community i be given by

$$Y_{1i}(\theta_i, \eta_i) = N_i \pi_i y_1(\theta_i, \eta_i), G_{1i}(\theta_i, \eta_i) = \theta_i Y_{1i}(\theta_i, \eta_i). \quad (5)$$

These are strictly increasing in local need and local cost. Hence implementation of the first-best requires information concerning these parameters.

¹⁰The analysis can easily incorporate an additional welfare weight on the utilities of the poor, by inflating v_p suitably: this would not alter the analysis qualitatively.

We now introduce some constrained versions of welfare-maximizing programs that will turn out to be useful in characterizing the outcomes of the centralized and decentralized systems along a number of different dimensions. These are therefore useful not for their own (i.e., normative) sake, but for describing the outcomes of different regimes. The *second-best allocation* results when the public program fails to incorporate information concerning the local needs and local costs in apportioning expenditures across communities. The expenditure level G_i for community i is then independent of θ_i and η_i . Assume however that the program is cost-effective and targeted perfectly within each community. In this case each poor person gets $y_{2i}(\theta_i, G_i) = \frac{G_i}{N_i \pi_i \theta_i}$ units of the good. The optimal expenditure grant G_{2i} for the program in community i solves:

$$\mathbf{E}_{\theta_i, \eta_i} \left[\frac{1}{\theta_i} v'_p \left(\frac{G_{2i}}{N_i \pi_i \theta_i}, \eta_i \right) \right] = (1 + \lambda). \quad (6)$$

The welfare gap between the first-best and second-best corresponds to the problem of lack of *community insurance*.

The *third-best allocation* corresponds to the situation where the the program is not cost-effective, so the good is procured at the maximal cost rate $\bar{\theta}_i$ instead of θ_i . This allocation satisfies the condition

$$\mathbf{E}_{\eta_i} \left[\frac{1}{\bar{\theta}_i} v'_p \left(\frac{G_{3i}}{N_i \pi_i \bar{\theta}_i}, \eta_i \right) \right] = (1 + \lambda) \quad (7)$$

and is smaller than the second-best level. The welfare gap relative to the second-best corresponds to the problem of *cost-effectiveness*.

The *fourth-best allocation* corresponds to the context where the government additionally fails to target delivery to the poor within any community. Then the good will be provided uniformly to the poor as well as non-poor within each community. The corresponding optimal expenditure grant G_{4i} solves

$$\mathbf{E}_{\eta_i} \left[\frac{1}{\theta_i} v'_p \left(\frac{G_{4i}}{N_i \pi_i \theta_i}, \eta_i \right) \right] = \left(1 + \frac{\lambda}{\pi_i} \right) \quad (8)$$

under the simplifying assumption that $y_n^*(\bar{\theta}_i, \underline{\eta}_i)$ is large enough relative to λ to ensure that the resulting scale of per capita public provision is smaller than market demand of the non-poor in all states.¹¹ The leakage of the good to the nonpoor is reflected in a higher markup

¹¹Under this condition, public provision displaces market expenditures of the nonpoor, so the marginal

factor used for the fiscal cost of the program ($\frac{\lambda}{\pi_i}$ instead of λ). This causes public supply to shrink further, by an extent that depends on the poverty rate. The welfare gap between the fourth-best and third-best allocations thus reflects the problem of *intra-regional targeting*.

In the fourth-best allocation, per capita grants are higher for communities with a higher poverty rate, to make allowance for the lower rate of intra-regional leakages. A converse problem of *inter-regional targeting* arises when per capita grants do not adjust for differences in poverty rate, so that poorer regions are characterized by lower per capita grants. In such situations, the poor who happen to be located in communities with a higher poverty rate would end up worse off than the poor in more prosperous communities. There is no reason for such a pattern to arise in a constrained welfare maximizing exercise, but we shall see that political economy factors may occasionally give rise to such problems below.

3 Political Capture

This section explains the actual objectives of governments at local and national levels, in particular the causes and extent to which they are prone to anti-poor bias owing to asymmetries in political awareness and participation of the poor in special interest politics. We draw on the analysis of this issue in our related papers (Bardhan and Mookherjee (1999, 2000)), which draws upon the earlier work of Baron (1994) and Grossman and Helpman (1996).

Consider first the nature of policy preferences of local governments, elected following ‘Downsian’ competition between two parties to win a majority of votes from local residents. Assume that a given fraction α_p of the poor in every district are ‘informed’ voters, who cast their vote partly on the basis of their utility difference from the policies espoused by the two parties. A higher fraction α_n of the non-poor in all communities are informed voters. This reflects higher levels of literacy and political awareness of the non-poor. Evidence in favor of this hypothesis for the US is presented by Delli Carpini and Keeter (1996, Ch.

value to the latter equals the market price θ_i , which is smaller than the fiscal cost of the provision. Moreover, the marginal value of the good provided to the poor exceeds θ_i .

4).¹² While we are not aware of any evidence concerning variations in political awareness with respect to income or education status in developing countries, it seems plausible that acquiring and processing information is costly and time-consuming for the poor who are resource-constrained, illiterate, and overworked.¹³

Another source of asymmetry is in participation in special interest group activity, such as election campaigns. Election campaigns are useful to sway the votes of uninformed voters, and special interest lobbies can seek to purchase ‘influence’ by contributing to these campaign funds. In the context of the United States, there is evidence covering the period 1952-88 that the propensity to contribute money, attend meetings and work on campaigns increased sharply with family income (Rosenstone and Hansen (1993, Table 8-2). We capture the asymmetry between poor and nonpoor groups simply by assuming that only the nonpoor form a lobby that contributes to the election finances of the two parties in exchange for special favors.

In the Downsian tradition, we assume that each party can credibly commit to a policy platform. Such commitments could be the outcome of long-term reputational concerns of the two parties. The objective of each party is thus to maximize the probability of winning the election, while that of the lobby is to maximize the utility of a representative non-poor person.

Voting behavior is ‘probabilistic’: it additionally depends on intrinsic party loyalties of voters. The two parties are distinguished by an intrinsic advantage awarded to one party on some ideological or non-economic dimension, or on the basis of past performance. These ‘ideological’ preferences are dispersed throughout the population according to a given uniform distribution. They may also be subject to random shocks that are the sum of a district-specific ‘swing’, and a nationwide ‘swing’ common to all districts. Informed voters

¹²For instance, the most informed citizens were older males whose family incomes exceeded \$50,000 annually. High income groups are overrepresented in the top 20% informed group relative to their demographic weight by 25%, and middle income groups by 18%. Low income groups in contrast were underrepresented in this group by 44%. These information gaps have remained stable over the last thirty to forty years.

¹³Indian politics is continuously marked by a reference to ‘poor vote banks’, in which poor uninformed voters are herded by community leaders and ‘delivered’ to electoral candidates.

then trade off their party loyalties against the utility difference resulting from the electoral platforms of the two parties. Uninformed voters trade-off their loyalties against the difference in campaign spending. The randomness in voter loyalties implies that neither parties nor lobbies can predict which party will eventually win the election. However, one party is *ex ante* favored to win the election, if the two parties select the same policy platforms. This asymmetry implies that the two parties do not select the same policy platforms in equilibrium.

It can be shown that in equilibrium each party selects a policy to maximize a linear welfare function, where the favored party assigns a greater welfare weight to the utility of the non-poor than the handicapped party, in exchange for higher campaign contributions from the nonpoor lobby. Specifically, party $k = 1, 2$ in elections to local government in district i selects a policy to maximize a welfare function

$$\alpha_p \pi_i U_{pi} + (\alpha_n + \chi_i \phi_i^k)(1 - \pi_i) U_{ni} \quad (9)$$

where U_{pi} and U_{ni} respectively represent the utilities of the poor and nonpoor in district i , ϕ_i^k is the probability that party k wins the election, and χ_i is a parameter which reflects the relative importance of the uninformed voters in this district and how sensitive they are to campaign spending. Apart from the effect of lobbying, an additional source of anti-poor bias in policy results from the lower level of political awareness among the poor. Let $\delta_i^{lk} \equiv \frac{\alpha_n + \chi_i \phi_i^k}{\alpha_p}$ denote the resulting premium placed on the welfare of the nonpoor, relative to the poor person: this parameter represents the extent to which the party in power in the local government of district i is captured. One would therefore expect that everything else remaining the same, the extent of capture will be increasing with respect to the local poverty rate, since this will tend to increase the relative importance of uninformed voters (hence the value of campaign finance to the parties), as well as the gap in political awareness between the poor and the nonpoor.

Turn now to the national level. Consider the case where each district first selects a representative to a national assembly on the basis of a majority vote election. The national government is then composed of the party winning a majority of seats in the national assembly. The same results obtain when there is a single nation-wide voting district for the

national election, or in a Presidential system where the nation's President is selected on the basis of winning a majority of first-past-the-post elections from different districts. Under the assumption that the number of districts is large, and that district-specific swings are also uniformly distributed, it can be shown that the objectives of the two parties in a national election aggregates their objective in local elections across different districts, resulting in a linear welfare function:

$$\Sigma_i[\alpha_p \pi_i U_{pi} + (\alpha_n + \chi \phi^k)(1 - \pi_i)U_{ni}] \quad (10)$$

where χ is a parameter determining the effectiveness of campaign funds in mobilizing the support of uninformed voters at the national level, and ϕ^k the probability that party k wins the election at the national level. The resulting level of capture (i.e., welfare weight on a nonpoor person *vis-a-vis* a poor person) in national policy-making of party k is thus $\delta^{nk} \equiv \frac{\alpha_n + \chi \phi^k}{\alpha_p}$.

It is clear that the degree of relative capture at the two levels will depend on factors determining the relative effectiveness of campaign finance, and the relative win probabilities for each party in local and national elections. In Bardhan and Mookherjee (1999, 2000) it is shown that these depend in turn on the following political and demographic variables: (i) relative intensity of electoral competition at the two levels (which depends on the heterogeneity of voter party loyalties across districts), (ii) relative cohesiveness of the interest group and of voter awareness at national and local elections (which depends in turn on the nature of the media and extent to which the nonpoor is politically fragmented on other dimensions), and (iii) heterogeneity of different districts with respect to their poverty rates. There is no *a priori* reason to expect greater or lower capture at the local level on average, as these different factors can pull in different directions. The situation is further complicated when the national government is elected on the basis of proportional representation. The most common pattern that typically emerges from the theory is one where the extent of capture of local governments in different districts varies widely, with the level of national capture occupying an intermediate value. To illustrate the intuition underlying this, if districts vary with respect to their poverty headcount rates, poorer districts will tend to have a larger fraction of uninformed voters, and local governments in those districts will

therefore be prone to greater capture. Since the national poverty rate will be an average of the district-level poverty rates, so will the national fraction of uninformed voters (hence national capture) tend to average the local fractions of uninformed voters (local capture).

From our perspective, however, the extent of political capture at any level of government may be treated as an exogenous parameter, which depends on a variety of political institutions. No prior restriction can be imposed, except perhaps for the pattern wherein poorer districts are prone to greater capture at the local level, with national capture somewhere in-between the extremes of the range of local capture.

4 The Centralized Regime

Consider first the centralized system in which responsibility for delivery is delegated to bureaucrats appointed by the central government. We assume a simple structure of the bureaucracy where a bureaucrat is appointed for any given district. The role of the bureaucrat is to monitor local cost and need, procure the good and then identify who the good should be delivered to. Information concerning these local parameters is costlessly available *ex post* to the assigned bureaucrat. In contrast the central government does not have access to any effective system of supervision of the information and actions chosen by the bureaucrat, who is remunerated a fixed salary.

If the bureaucrats were honest, the system would work as follows. Let B_i denote the local bureaucrat assigned to community i , whose role is to gather information about η_i and θ_i , and send this information to the central government (denoted CG). Having received this information, CG will authorize the first-best expenditure $G_{1i}(\theta_i, \eta_i)$ to B_i , who would then use this to locally procure supplies of the good at cost θ_i , and then distribute the good to the poor. If B_i were honest, he would report local information to CG honestly, procure the good at true cost, and distribute the good to the local poor.

Suppose now that B_i 's values swing to the opposite end of the ethical spectrum, and seeks to maximize personal income instead. Assume that the bureaucrat has lexicographic preferences: maximization of personal income is the primary objective, but if two courses of

action generate the same level of personal income then he prefers to do what he is supposed to do: help the poor.

Having collected information concerning η_i, θ_i and who is poor within each community, would it then be in B_i 's interest to faithfully implement the first-best allocation? B_i has discretion over reports to submit to CG concerning local conditions, and then over how to spend the allotted budget. We assume that government auditors can verify the physical procurement and delivery level Y_i in each community i , but not the realization of local need and costs. So B_i can claim a certain market price of this good : $\tilde{\theta}_i \in [\theta_i, \bar{\theta}_i]$, which overstates the true cost, and pocket the difference.¹⁴

B_i can go further and earn illegal income by diverting part of the supply to community i to the market, instead of delivering it to the poor. Given the nature of the good, B_i cannot utilize any form of nonlinear pricing on the 'black' market: resale is costless, and the discrepancy between discriminatory prices would be eliminated by arbitrage. It is clear that B_i can sell (at a price of θ_i) upto a total quantity of $N_i(1 - \pi_i)y_n^*(\theta_i, \eta_i)$ units on the black market, i.e., the total market demand of the nonpoor in this community. In case the available quantity Y_i exceeds this amount, it would not pay B_i to attempt to sell more at a lower price, since the market demand of the nonpoor is price-inelastic. In the event that $Y_i > N_i(1 - \pi_i)y_n^*(\theta_i, \eta_i)$, the optimal course of action for the corrupt bureaucrat therefore is to sell exactly $N_i(1 - \pi_i)y_n^*(\theta_i, \eta_i)$ units on the black market, and (given the assumption of lexicographic preferences) deliver the rest to the poor. In any given state, therefore, the poor do receive some of the good allocated to that community, but only if the aggregate supply is large enough relative to the market demand of the nonpoor.

An alternative reason that bureaucrats might deliver services to the poor to some extent, of course, is that some degree of monitoring of their activities does take place, either by an

¹⁴It is plausible that information concerning physical procurement and delivery levels is 'hard' information, verifiable by third parties, while information concerning local cost and need is nonverifiable. Government record-keeping frequently concerns physical purchases and movement of goods, rather than their temporal market or consumption valuations. We are also assuming that information concerning community supplies Y_i is verified by government auditors, but not utilized by CG to design superior performance based incentive schemes for bureaucrats.

occasionally honest (and courageous) supervisor, subordinate, or other whistleblowers (such as the media), and verifiable reports of corruption are followed by imposition of penalties by the central government. The resulting compliance incentive would then be similar to that of corrupt taxpayers who comply partially with their tax obligations owing to a certain probability that their evasion activities will be detected and punished, as in the theory of Allingham and Sandmo (1972) or Yitzhaki (1974). Then a positive fraction of supplies will be targeted to the poor. We shall later show that our main qualitative results would be unaltered in this extension. For the time being we therefore abstract from control of bureaucratic corruption achieved by monitoring.

With reported expenditure $\sum_i \tilde{\theta}_i Y_i$ being reimbursed by the government, the illegal earnings of B_i equals

$$\Pi_i(\tilde{\theta}_i, Y_i | \theta_i, \eta_i) \equiv (\tilde{\theta}_i - \theta_i)Y_i + \theta_i \min\{Y_i, N_i(1 - \pi_i)y_n^*(\theta_i, \eta_i)\}, \quad (11)$$

i.e., the sum of income realized from cost-padding of the supply Y_i , and from black market sales to the nonpoor. Then B_i will optimally want to expand the scale of the program and pad costs to the maximum permissible extent. Realizing that the mechanism gives rise to such incentives, and that only a fraction of expenditures actually reach the poor, the central government in turn will seek to restrict expenditure allotments, even if it is not subject to any form of ‘capture’ by special interest groups (e.g., of the nonpoor that strive to limit the taxes that finance the program).

It is now possible to derive the allocation that will actually result when delivery is delegated to corrupt centralized bureaucrats. From the discussion above, it is evident that such a bureaucrat will allocate a given supply Y_i of the good on a preferred basis to the nonpoor upto their total market demand, and thereafter allocate the remainder to the poor. In a given state (θ_i, η_i) , then the net utility gains of the two groups respectively are

$$U_n(Y_i; \theta_i, \eta_i) = 0 \quad (12)$$

and

$$U_p(Y_i; \theta_i, \eta_i) = v_p(\max[\frac{Y_i - N_i(1 - \pi_i)y_n^*(\theta_i, \eta_i)}{N_i\pi_i}, 0], \eta_i) \quad (13)$$

The nonpoor derive no advantage whatsoever, since they substitute open market purchases by black market purchases at the same price. The poor do get some of the good delivered in states when there is something left over after meeting the black market demand of the nonpoor. The entire leakage is pocketed by the bureaucrat.

Suppose B_i is granted an expenditure limit of G_i for market purchases of the good by the central government. Then B_i has to decide how to use this budget: i.e., must select how much of the good Y_i to purchase, and what cost $\tilde{\theta}_i$ to report on the accounts, subject to the constraint that the expenditure on the books $\tilde{\theta}_i Y_i$ does not exceed G_i , and $\tilde{\theta}_i$ does not exceed the maximum possible cost $\bar{\theta}_i$. His problem is to maximize his total income resulting from these decisions, as given by (11), subject to these two constraints. The solution to this problem is to set $\tilde{\theta}_i = \bar{\theta}_i$, and $Y_i = \frac{G_i}{\bar{\theta}_i}$, i.e., pad accounting cost and expenditure of the budget to the maximum permissible limit.¹⁵

The maximized income of B_i then equals

$$\Pi_i(G_i|\theta_i, \eta_i) = \min[G_i, G_i[1 - \frac{\theta_i}{\bar{\theta}_i}] + \theta_i(N_i - P_i)y_n^*(\theta_i, \eta_i)] \quad (14)$$

which is a strictly increasing, concave function of allotted grant G_i . For small grant levels, B_i can pocket the entire allotment himself, since the actual procurement cost can be recovered completely from black market sales, so that the income of B_i equals the reported expenditure level reimbursed by the government. This is always true for instance for grants smaller than $\underline{G} \equiv \underline{\theta}_i y^*(\underline{\theta}_i, \eta_i)$, no matter what state of the world prevails in community i . Such expenditure allotments are then entirely diverted to the black market. For grants higher than \underline{G} , black market demand in some states falls below the level of good procured, and the surplus is distributed to the poor. When such states occur, B_i 's earnings are smaller on account of his inability to divert the entire supply to the black market. But he continues to earn fraction $1 - \frac{\theta_i}{\bar{\theta}_i}$ of the grant owing to the cost-padding. This income provides the

¹⁵The optimality of reporting $\tilde{\theta}_i = \bar{\theta}_i$ is evident from (11) no matter what the supply decision Y_i is: it maximizes the income from cost-padding (the first term on the right-hand-side of (11)). And given $\tilde{\theta}_i = \bar{\theta}_i$, it is always optimal to increase Y_i as far as the expenditure budget will allow, since this increases cost-padding income by more ($\bar{\theta}_i$ per unit) than it does procurement cost (θ_i per unit), so even if B_i throws away the incremental supply of the good his income will be higher.

underlying incentive to expand procurement as far as is permitted by the budget, and is the basic reason that the poor end up with something at all.

If bureaucrats are employed on a fixed salary basis, it follows from (14) that there is no way for the local bureaucrat to credibly communicate local cost or needs to the central government. Any attempt by the latter to elicit this information via a reporting or budgeting system would cause local bureaucrats to manipulate these reports to maximize their allocated budgets. Hence actual expenditure allocations made by the central government cannot be conditioned on information about local cost or need. Community grants can be based only on publicly available information, such as relative population sizes and poverty rates. The centralized system then fails to achieve any community insurance.¹⁶

In addition, the allocation fails to be cost-effective, and realizes a poor standard of targeting. Leakages are entirely intramarginal: above a threshold (determined by the market demand of the nonpoor) increased expenditure allotments are targeted one hundred percent to the poor. And only the poor benefit from the existence of such a program, whilst the nonpoor end up paying the taxes that finance it. The only rationale for the central government to maintain the program is the political support of the poor.

4.1 Central Grants

The outcome of the centralized regime eventually depends on the grants G_1, \dots, G_m authorized for different regions by the central government. Assuming that bureaucrats form an insignificant proportion of the voting population, the policy platform of party $k = A, B$ in

¹⁶It will achieve some degree of intercommunity targeting if incentive mechanisms for bureaucrats can be utilized in some form. These could conceivably take the form of explicit performance incentive mechanisms which make compensation contingent on budget size or the aggregate scale Y_i of services delivered. Or they could result from competition for rents between upper and lower level bureaucrats.

national elections will seek to maximize¹⁷

$$W_k^C(G_1, \dots, G_m) = \sum_{i=1}^m \pi_i N_i \mathbf{E}_{\theta_i, \eta_i} v_p(\max(0, \frac{G_i}{N_i \pi_i \theta_i} - \frac{(1-\pi_i)}{\pi_i} y_n^*(\theta_i, \eta_i)), \eta_i) - \delta^{nk} (1 + \lambda) \sum_{i=1}^m G_i \quad (15)$$

The identity of the political party k enters only through the dependence of the capture coefficient δ^{nk} . So to simplify notation we shall henceforth suppress k in the notation. (15) shows how central grant allocations trade off between the effect of increased expenditure allotments on the preferences of poor informed voters (what may be referred to as the *vote-bank effect*, the first term on the right-hand-side of (15)), with the cost imposed on the nonpoor who fund the program through their taxes. Note that there is a nonconcavity with respect to the vote-bank effect, since only expenditures above a minimum scale actually reach the poor. Hence any expenditure allotments that are actually made, must be at a certain minimum scale, and do trickle down to the poor (at least in some states of the world). *From the standpoint of utilitarian social welfare, however, the outcome of a centralized bureaucracy is worse than the fourth-best allocation:* it fails to achieve any cost-effectiveness or community insurance, and targeting is poorer than uniform provision of the good within each community (the standard assumed in the fourth-best). We summarize these results in the following proposition (in which we use the term capture parameter to denote the relative welfare weight assigned to the nonpoor by the political party forming a government at the center).

Proposition 1 *The outcome of centralization is the following:*

¹⁷In deriving (15) we have made use of expressions (12, 13) above for the benefits U_p, U_n obtained by different classes from a given expenditure level, and the fact that the entire fiscal burden of the program is borne by the nonpoor, since they are the only ones paying any taxes. This expression also supposes that political parties at the national level are uncertain about the realization of local cost and need parameters in any given community. Alternatively, expenditure allotments are decided and held fixed for a given period of time, such as the period of incumbency following any election, during which local community costs and needs vary randomly. As explained above, the centralized system does not allow expenditure assignments to be conditioned on the realization of these shocks. Then voters' policy preferences will be based on the expected utility they expect to realize from any given expenditure assignments.

(i) The expenditure grant G_i allocated to community i by a central government with a capture parameter δ^n is independent of the actual realization of θ_i, η_i , and selected to maximize

$$N_i \pi_i \mathbf{E}_{\theta_i, \eta_i} [v_p(\max\{0, \frac{G_i}{N_i \pi_i \theta_i} - \frac{(1 - \pi_i)}{\pi_i} y_n^*(\theta_i, \eta_i)\}, \eta_i)] - \delta^n (1 + \lambda) G_i. \quad (16)$$

(ii) Given expenditure grant G_i , the good is procured at an ‘official’ price of $\bar{\theta}_i$, and the total supply of the good to community i is $\frac{G_i}{\bar{\theta}_i}$. Of this, the total amount delivered to the poor equals $\max\{0, \frac{G_i}{N_i \bar{\theta}_i} - N_i (1 - \pi_i) y_n^*(\bar{\theta}_i, \eta_i)\}$, the remainder being diverted to the black market.

(iii) The resulting level of (per capita) utilitarian welfare in community i

$$\pi_i \mathbf{E}_{\theta_i, \eta_i} [v_p(\max\{0, \frac{G_i}{N_i \pi_i \theta_i} - \frac{(1 - \pi_i)}{\pi_i} y_n^*(\theta_i, \eta_i)\}, \eta_i)] - (1 + \lambda) G_i \quad (17)$$

falls below the fourth-best level.

5 Decentralization

Now suppose that the responsibility for local procurement and allocation of the good is devolved to elected local governments, who are financed by grants from the central government. Taxing power as well as authority over intercommunity transfers is retained by the central government. The only switch relative to the centralized mechanism is that local authority over procurement and intracommunity allocation is shifted from a local bureaucrat who seeks to maximize illegal income, to a local government that represents the interests of local citizens. The local government has access to the same information concerning local cost and need parameters as centralized bureaucrats assigned to that district. Hence local governments are not subject to comparable agency problems of bureaucratic corruption that the central government is. This is the prime advantage of the decentralized program. To highlight it in its most pristine form we shall suppose that officials of the local government either procure and distribute the good themselves. Alternatively they delegate it to local bureaucrats that they appoint and closely supervise, thereby avoiding bureaucracy

agency problems. This corresponds to Robert Wade's depiction of the Korean irrigation bureaucracy.

The potential disadvantage of the decentralized system, however, is that it is subject to political capture: local governments will be subjected to pressures from the local nonpoor to divert the good to them instead. Intra-district targeting will be determined by the politics of local capture, while grant allocations to different districts will continue to be determined by the central government (given their expectations concerning targeting at the local level).

Consider first what happens within a given community i , after it has received grant G_i from the center. In the case where the grant is tied to purchases of the specific good in question, it is apparent that the local government will want to maximize the total supply of the good that it has available to distribute among its residents. Hence it will seek to procure the good from the market at the going price θ_i , and will derive no advantage from cost-padding (since any incomes earned thereby from cost-padding cannot be diverted to other goods favored by the nonpoor, given that the grant is tied).¹⁸ Hence the good will be procured cost-effectively under decentralization, allowing delivery cost economies relative to the centralized bureaucracy outcome. Such outcomes have frequently been noted in the Indian state of West Bengal following experiments with expenditure decentralization (Dasgupta (1997)). The principal reason is that decisions are made in the interests of users of the good, rather than of outside bureaucrats who derive no personal benefit from service delivery.

Given grant G_i , the local government will be able to procure $\frac{G_i}{\theta_i}$ units of the good. How will it allocate this supply among its residents? To study this, we need to calculate the net

¹⁸A similar conclusion holds (for most states of the world) even if the grant is not tied. Whatever benefits the nonpoor can derive from diverting the funds to their pockets or to other goods they favor, can equivalently be achieved by procuring this good and diverting the allocation of the good in their own favor (assuming that the grant in question does not exceed the value of the total market demand of the good of the nonpoor, which will turn out to be the case in equilibrium). This will be exactly correct if there is no uncertainty concerning local costs or needs. If there is some uncertainty, this condition may fail in certain states where local costs are high and needs are low, so that the market demand of the nonpoor falls below the grant. In those states, and only those, would there be pressure from the nonpoor to divert funds if the grant is not tied, e.g., by cost-padding or spending on other goods that they value more.

benefits of the two classes from receiving the good. These are given by

$$V_n(y_n, \eta_i) = \begin{cases} \theta_i y_n & \text{if } y_n \leq y_n^*(\theta_i, \eta_i) \\ \theta_i y_n^*(\theta_i, \eta_i) + v_n(y_n, \eta_i) - v_n(y_n^*(\theta_i, \eta_i), \eta_i) & \text{otherwise} \end{cases} \quad (18)$$

and

$$V_p(y_p, \eta_i) = \begin{cases} \theta_i v_p(y_p, \eta_i) & \text{if } y_p \leq y_p^*(\theta_i, \eta_i) \\ \theta_i v_p(y_p^*(\theta_i, \eta_i), \eta_i) + \theta_i [y_p - y_p^*(\theta_i, \eta_i)] & \text{otherwise} \end{cases} \quad (19)$$

where the expression for V_p assumes that the aggregate supply of the good is smaller than the first-best level.¹⁹ For the non-poor upto the point that supply is less than their market demand, the marginal value of the good exactly equals its market price θ_i , since public allocation will exactly substitute for market purchases. If they are supplied more than their market demand, they will have no option but to consume the entire supply, since they will be unable to sell it on the market: the poor will be unable to pay, while there is no scope for trades among the nonpoor. Then the marginal utility of the public allocation for the nonpoor drops below the market price θ_i . For the poor, on the other hand, the marginal value of the good will equal the marginal utility of consumption of the good, upto the point that they are allocated less than their own (income-unconstrained) market demand $y_p^*(\theta_i, \eta_i)$. Over this range, they will value it at more than the going market price θ_i . If they are allocated more than their market demand, they will seek to sell the good on the market. Given our assumption concerning the overall scale of supply to the community, if the poor are supplied more than their first-best allocation, the nonpoor must be supplied less than their first-best allocation (equal to their market purchase). Hence there will be (black) market demand of the nonpoor that the poor can offload their excess supplies to, at the market price θ_i . Over this range the poor's marginal valuation of the good equals the market price. The nature of marginal valuations of the good to the two groups are illustrated in Figure 1.

INSERT FIGURE 1 HERE

¹⁹It is possible, of course, that this condition is not satisfied in equilibrium for some states . Then the value of the good to the nonpoor will take a more complicated form, but will not alter our qualitative results.

Given the political compulsions on the local government, represented by capture parameter δ_i^l , it will allocate total community supply Y_i between the two classes in state (θ_i, η_i) to maximize the welfare function (again suppressing party identity k in the notation):

$$\pi_i V_p(y_p, \eta_i) + \delta_i^l (1 - \pi_i) V_n(y_n, \eta_i) \quad (20)$$

subject to the resource constraint

$$\pi_i y_p + (1 - \pi_i) y_n \leq \frac{Y_i}{N_i}. \quad (21)$$

The effect of local capture is to inflate the marginal value of delivery to the nonpoor by the capture coefficient δ_i^l . Any given unit of the good will be allocated upon comparing this inflated marginal value for the nonpoor with the marginal value of the good to the poor. Define $y^*(\delta_i^l, \eta_i, \theta_i)$ by the solution to

$$v'_p(y^*, \eta_i) = \delta_i^l \theta_i \quad (22)$$

if it exists, and 0 otherwise. Then the poor will be served on a preferential basis upto consumption level of y^* , following which the nonpoor will be served upto their market demand, with sharing between the two groups thereafter. The pattern of targeting that results is illustrated in Figure 2.

INSERT FIGURE 2 HERE

The decentralized system thus targets small grants better than the centralized one, provided local capture is not so great that it prevents any delivery to the poor altogether. The decentralized system initially serves poor residents on a preferential basis, after which it switches to the non-poor, exactly reversing the sequence in a centralized system. For large grants, on the other hand, the centralized mechanism could end up delivering more to the poor: once the black market demand of the nonpoor has been met, the poor get all of the remainder. *Hence the relative success of the two systems depends on the scale of expenditures on the program: small grants will be targeted better in the decentralized regime, and larger grants will be targeted worse.* This result will continue to hold even in other settings where the centralized system ensures some degree of targeting owing to monitoring

of bureaucrats: as noted above such a system would generally induce a positive fraction (less than unity) of all expenditures to be delivered to the poor.

Finally we need to consider how the center determines expenditure grants in the decentralized environment. This raises a number of issues concerning how local and central governments negotiate with one another. For instance, if the same party governs both at the local and central level, it is plausible that the governments at two levels will cooperate with one another, sharing information and working out a common strategy that advances the general interest of the party. But conflict may result if different parties govern at the two levels. Since our main objective is to compare the outcome with that of a centralized bureaucracy, we shall consider the scenario which compares exactly with our formulation of the latter mechanism. Specifically, assume that each local government behaves opportunistically, and does not cooperate with the central government. It therefore seeks to obtain as large a grant as possible for itself from the center, and misrepresents its information about local need and cost in order to manipulate the grant it receives. Moreover, the party in power at the central government decides community grants to serve its political motives at the national level. As a result, community grants in the decentralized regime fail to achieve any community insurance, and end up being independent of local shocks. To the extent that the assumption of total noncooperation is not realistic, the local government mechanism may be able to achieve some degree of community insurance. In that case the balance will tilt further in favor of the decentralized mechanism.

Let $y_k(y; \delta_i^l, \eta_i, \theta_i), k = p, n$ denote the intracommunity allocation generated by a per capita supply of y to the community, when the local capture coefficient is δ_i^l , and the realized local need and cost parameters are (η_i, θ_i) . Then a central government with capture parameter δ^n will determine the allocation of expenditure grants across communities to maximize

$$\begin{aligned}
 W_k^L(G_1, \dots, G_m) = & \sum_{i=1}^m N_i \{ \pi_i \mathbf{E}_{\theta_i, \eta_i} [V_p(y_p(\frac{G_i}{N_i \delta_i^l}; \delta_i^l, \eta_i, \theta_i)) \\
 & + \delta^n (1 - \pi_i) V_n(y_n(\frac{G_i}{N_i \delta_i^l}; \delta_i^l, \eta_i, \theta_i)) \\
 & - \delta^n (1 + \lambda) G_i \}. \tag{23}
 \end{aligned}$$

The first term on the right-hand-side of (23) represents the effect on votes from the poor.

The second represents the political benefit imputed to delivering the good to the nonpoor, while the third incorporates loss of support from the nonpoor since their tax payments eventually fund the system. The last effect clearly dominates the second, since the nonpoor pay for the program but receive only part of the benefit. Overall, the program causes political parties to lose support from the nonpoor, who will exercise pressure to limit its size.

Proposition 2 *The allocation resulting under expenditure decentralization is the following.*

(i) *Community grant G_i for community i is independent of local shocks, and is chosen to maximize*

$$\pi_i \mathbf{E}_{\theta_i, \eta_i} [V_p(y_p(\frac{G_i}{N_i \theta_i}; \delta_i^l, \eta_i, \theta_i), \eta_i) + \delta^n (1 - \pi_i) V_n(y_n(\frac{G_i}{N_i \theta_i}; \delta_i^l, \eta_i, \theta_i), \eta_i)] - \delta^n (1 + \lambda) G_i \quad (24)$$

(ii) *Given grant G_i , the good is procured at true cost θ_i . It is thereafter allocated as follows: it is allocated entirely to the poor until a per capita procurement of $\pi_i y^*$, above which the good is supplied entirely to the nonpoor upto their market demand y_n^* , with any excess shared between the two groups so as to equate V_p' and $\delta_i^l V_n'$.*

(iii) *The resulting level of (per capita) social welfare in community i*

$$\pi_i \mathbf{E}_{\theta_i, \eta_i} [V_p(y_p(\frac{G_i}{N_i \theta_i}; \delta_i^l, \eta_i, \theta_i), \eta_i) + (1 - \pi_i) V_n(y_n(\frac{G_i}{N_i \theta_i}; \delta_i^l, \eta_i, \theta_i), \eta_i)] - (1 + \lambda) G_i \quad (25)$$

equals the second-best welfare level if there is no capture at either central or local levels. If the marginal utility of the good to the nonpoor is positive for any level of consumption, then welfare in community i under the decentralized system tends to 0 if the local capture coefficient δ_i^l becomes arbitrarily large.

Corollary to Proposition 2 *The decentralized mode welfare dominates the centralized mode in the absence of any political capture at either central or local levels. The centralized mode dominates instead if capture at the local level is sufficiently large relative to that at the central level.*

The relative levels of political capture at central and local levels thus plays a key role in determining which of the two modes generates higher welfare. When there is no capture at either level, the decentralized mode achieves second-best welfare (i.e., perfect cost-effectiveness and targeting), in contrast to the centralized mode (which performs no better than the fourth-best allocation). In this case decentralization enables a reduction in the bureaucratic corruption that afflicts the centralized system, without creating any countervailing distortions.

At the other extreme as the extent of local capture becomes arbitrarily large, decentralization serves to divert the good almost entirely to the nonpoor (as long as they always derive positive marginal utility from consuming the good). Knowing this, the center (irrespective of the level of capture at the national level) does not find it worthwhile to allocate any expenditure grants to local communities, since the value of the good to the nonpoor falls below the tax burdens they bear to support the program; the program tends to close down altogether. If the centralized system provides positive community grants, it must then achieve a superior welfare level, owing to its ability to deliver to the poor in some (albeit imperfect) degree. Hence in circumstances where the problem of local capture is sufficiently great relative to that at the national level, centralization may be the lesser of the two evils.

In one special case, however, the two regimes can be unambiguously ranked, where the demand of the non-poor is deterministic. In the contexts of certain commodities, such as schooling or health care, it is plausible to suppose that the market demand of the nonpoor does not vary with the state, since the marginal utility of the service beyond a certain level is zero. Ignoring differences in quality, it is frequently the case that the rich send all their children to school and purchase all the health care they need on private markets, irrespective of what market prices happen to be (i.e., over the range of normal variation of market prices), and would derive no further utility from additional levels of provision. If the local community is sufficiently polarized, the nonpoor consist entirely of such rich households. In such a case, once the market demand of the nonpoor has been met, local governments will not allocate any further units of the good to them, irrespective of the level of local capture. Then intracommunity targeting in the decentralized system is unambiguously better than in the centralized system: see Figure 3 where the decentralized delivery

pattern is represented by the locus $abce$, and the centralized pattern by locus $adce$. It is therefore natural to expect in this case that the decentralized system *always* dominates the centralized one, irrespective of the nature of political distortions, or underlying uncertainties in local cost and need. We have been able to verify this in the case where the utility function of the poor takes the form $v_p(y, \eta_i) = \eta_i v(y)$, where $v(\cdot)$ is a homothetic function.

INSERT FIGURE 3 HERE

6 Targeting Comparisons

In order to gain additional insight into the relative targeting performance of the two delivery modes, it is useful to focus exclusively on this dimension, and abstract from problems of cost-effectiveness and community insurance. This is the situation where local cost and need parameters are not variable, or can be observed by the central government and used in the determination of fiscal grants. Then the two regimes can be compared entirely with respect to their performance with respect to inter- and intra-regional targeting.

In this case the market demand y_n^* of the nonpoor in any given region is known to the central government while setting the level of the expenditure grant G_i . In a centralized bureaucracy, the leakage to the nonpoor can be predicted perfectly by the central government. In what follows we can suppress the notation for the local cost or need parameter. Since the welfare function is additively separable across regions we can also focus on a given region i . The service delivery to the poor y_i^c induced by the centralized regime maximizes

$$\pi_i v_p(y) - (1 + \lambda) \delta^n \theta_i [\pi_i y + I(y)(1 - \pi_i) y_n^*] \quad (26)$$

where $I(y)$ denotes an indicator function taking the value 1 if $y > 0$, and 0 otherwise. Provision of the good to the poor thus entails a ‘fixed’ cost consisting of the leakage of the good to the nonpoor, besides the variable cost. It is evident that if the service is delivered in positive quantity to this region, y_i^c satisfies the first-order condition

$$v_p'(y_i^c) = (1 + \lambda) \delta^n \theta_i \quad (27)$$

The corresponding fiscal grant per poor resident in the region equals

$$\frac{G_i^c}{N_i \pi_i} = \left[\frac{1 - \pi_i}{\pi_i} y_n^* + y_i^c \right] \theta_i \quad (28)$$

in the case that the service is provided at all to this region, and zero otherwise. The resulting level of (per capita) welfare in this region is

$$w_i^c \equiv \pi_i v_p(y_i^c) - (1 + \lambda) \theta_i [\pi_i y_i^c + I(y_i^c)(1 - \pi_i) y_n^*]. \quad (29)$$

Figures 4 and 5 depict the pattern of service provision and fiscal grants across regions of varying poverty rates. The centralized delivery pattern is discontinuous, represented by the union of the two lines bc and def . The decentralized pattern is depicted by the locus $adeg$. In regions with a low poverty rate, the relative number of nonpoor residents is large, causing larger leakages to the black market, raising the cost of delivering a given level of services to a representative poor person. Hence for sufficiently prosperous regions, marked L in the figures with poverty rate below π_1^* , the centralized system will not provide any services at all. Regions above this threshold will see a uniform level of delivery to poor residents, as shown by the line def in Figure 4 (assuming that the cost is the same across all regions, otherwise corresponding adjustments will be made). The size of the grant necessary to achieve this has make allowance for the anticipated leakage to the nonpoor, and this leakage rate is lower in poorer regions. Hence for regions M and H above the threshold, the size of the per capita grant will be smaller in poorer regions, creating a misleading appearance of an inter-regional targeting failure (see Figure 5, where the per capita grant above the threshold is depicted by the locus df). The grant is smaller in poorer regions precisely because such regions are expected to experience lower (intra-regional) targeting failures, so a smaller grant is necessary to ensure a uniform level of service.

INSERT FIGURES 4 AND 5 HERE

Now turn to the decentralized system. Here too the nature of leakage to the nonpoor can be predicted by the central government, given its knowledge of the level of capture that the local government is prone to. The per capita supply y^d to the community may be set at less than $\pi_i y^*$, in which case it is targeted perfectly. Or it exceeds $\pi_i y^* + (1 - \pi_i) y_n^*$, in which

case the leakage to the nonpoor exceeds that under the centralized system. Intervening levels of supply are not worthwhile since all units at the margin go to the nonpoor, the benefit of which outweighs the fiscal cost, so the central government would cut the size of the grant to finance a per capita provision of $\pi_i y^*$. In countries with weak fiscal capacity and with large black markets (i.e., both λ and y_n^* are large), the scale of the program is not going to be large enough to accommodate the entire demand of the nonpoor. In such a case, the program will be set at a level not exceeding $\pi_i y^*$. Assume for the time being that this is indeed the case: it will turn out to not matter for our main conclusion below.

Then the decentralized system will achieve perfect intracommunity targeting, but is constrained in the scale of the program by the ceiling $\pi_i y^*$ that depends on the extent of local capture δ_i^l (depicted by the locus *heg* in Figure 4. As local capture increases this ceiling is lowered. The optimal provision of service to the poor y^d under decentralization solves the problem of maximizing

$$\pi_i[v_p(y) - \delta^n(1 + \lambda)\theta_i y] \quad (30)$$

subject to the constraint $y \leq y^*(\delta_i^l)$. In words, it is constrained both by political capture of the central government (since this is reflected in its imputed cost of financing the grant that is incorporated in the objective function (30)), and the capture of the local government (which determines the amount of provision that the local government can be trusted to target correctly, as incorporated in the constraint). The unconstrained solution coincides exactly with the optimal provision y_i^c in the centralized regime (whenever the latter is positive), because in such a case only the level of capture at the central level matters (and in identical fashion) in both the centralized and decentralized regimes. The resulting pattern of deliveries to the poor under decentralization is depicted by the locus *adeg* in Figure 4, and the corresponding per capita grants is depicted by the locus *aeg* in Figure 5.

Service delivery levels under the two regimes diverge in either of two situations. First, in especially poor regions marked H in the figures, the constraint concerning local targeting in problem (30) will bind, whence the decentralized system will deliver a smaller service level than the centralized one. In Figures 4 and 5 we assume that the threshold π_2^* where the constraint binds in the decentralized system exceeds the threshold π_1^* where the centralized

system begins to deliver to the poor, thus creating the three regions L, I and H.²⁰ It is evident that the local targeting constraint binds if and only if $\delta_i^l \geq (1 + \lambda)\delta^n$, so the threshold π_2^* corresponds to a level of local capture which equals $(1 + \lambda)\delta^n$. This gets reversed in regions with poverty below the threshold π_1^* , the decentralized system delivers to the poor, while the centralized one does not. Hence the levels of service delivery (and corresponding per capita welfare levels) under the two regimes cannot be ordered in general. In the low poverty L region the delivery performance of the decentralized system is superior, and so is the level of welfare generated. In contrast, the centralized system delivers more in the high poverty H region, though at the expense of a lower targeting efficiency, so the welfare comparison is ambiguous here. However for sufficiently large poverty rate the targeting inefficiency of the centralized system shrinks to zero, and so does the service delivery rate under decentralization. Ultimately the centralized system must therefore generate higher welfare. In the I region with intermediate poverty the two regimes deliver the same amounts to the poor, but the centralized system experiences targeting failures, so the decentralized regime welfare dominates here.²¹

The corresponding pattern of (per capita) fiscal grants in the two systems is depicted in Figure 5. Since the decentralized system involves no leakage at all, the per capita grants follow the same shape as the service delivery levels. In the low poverty L region, it therefore generates a uniform per capita grant, higher than under the centralized system. In all other

²⁰The reader can check that our qualitative conclusions below will continue to hold in the opposite case.

²¹The conclusion that decentralization dominates in terms of service delivery and welfare in low and intermediate poverty regions will be further reinforced if we drop the constraint the per capita delivery under the decentralized system will be set at or below $\pi_i y^*$, so as to ensure absence of any targeting failures. If the scale of the program is larger, thus permitting the decentralized system to divert part of the incremental services to the nonpoor, the service delivery levels to the poor must also increase. This is because the sole rationale for the expansion of the program from the perspective of the party in power at the central government is that the benefit in terms of votes of the poor outweigh the cost of the expansion that falls on the nonpoor. Welfare must go up since the social welfare function assigns a higher relative welfare weight to the interests of the poor than do political parties. At the other extreme, if the local poverty rate is sufficiently high, the targeting failures under decentralization will be large enough to ensure that service levels and welfare will drop below that under centralization. So the qualitative conclusion is unaffected.

regions, it is easy to see that the centralized regime involves higher per capita grants.²² Excluding the L region, therefore, decentralization will tend to shrink the volume of fiscal transfers, owing partly to better targeting, and partly to lower service deliveries in the H region.

Underlying the contrast between outcomes of the two regimes is the reversal of the pattern of targeting failures. The centralized regime experiences the greatest targeting failures in the low poverty regions, owing to large black markets. The decentralized regime experiences the greatest targeting failures in the high poverty regions, owing to poor governance at the local level. The poor in the high poverty regions tend to be worse off as a result of decentralization, while the poor in the low poverty regions benefit.

These results are qualitatively unaltered in the case where targeting incentives for central bureaucrats are provided by monitoring and punishment threats. Assume that with a given probability p the corrupt activities of a bureaucrat will be exposed and followed by imposition of penalties that are proportional to the amount of the service diverted to the black market. With probability $1 - p$ no penalties will be imposed. Then the bureaucrat's decision concerning diversion will be similar to the tax compliance decision studied by Allingham and Sandmo (1972) and Yitzhaki (1974). If the bureaucrat's relative risk aversion is constant, a constant fraction of services will be diverted to the black market, unless the bureaucrat is constrained by the size of the demand on the black market. Since the black market is smaller in a low poverty region, it is plausible that a lower fraction will be diverted to the black market in such regions. This generates exactly the same pattern of targeting failures as in the case without any monitoring: targeting failures are higher in low poverty regions, the opposite of the outcome under decentralization.

One difference, however, from our previous analysis is that there may no longer be a nonconvexity in centralized provision. If region i is characterized by a fraction c_i which is diverted to the black market, where c_i is a decreasing function of the poverty rate, then the service delivery level in the centralized mode for this region will solve the first order

²²To the right of π_1^* the centralized regime provides the same or greater service level to the poor, and also involves leakages to the nonpoor.

condition

$$(1 - c_i)v'_p(y_i^c) = \delta^n(1 + \lambda)\theta_i \quad (31)$$

and service provision to the poor is an increasing function of the poverty rate, just as before. The difference is that the program need not be shut down in regions with very low poverty rates. The corresponds patterns of service delivery and fiscal grants are depicted in Figures 6 and 7 respectively.²³ All the previous qualitative results continue to hold, with the exception that it is now possible that the size of grants under the decentralized regime is uniformly lower in all regions.

INSERTS FIGURES 6 AND 7 HERE

7 Concluding Comments

The purpose of this paper has been to develop an analytical framework to discuss the effects of delegating authority over the delivery of an antipoverty program to local governments rather than to bureaucrats appointed by a central government that is ill equipped to monitor and control discretionary behavior at the local level. Experimentation with such patterns of decentralization have recently become widespread throughout developing countries, motivated primarily to reduce corruption and targeting inefficiencies in the delivery system. The nature of decentralization, their effects and corresponding policy implications differ fundamentally from those typically analyzed in the traditional literature on fiscal federalism. Local governments are granted a very limited form of autonomy, only over procurement and distribution of the merit good in question, while central governments retain all authority over revenues and disbursement of grants to local areas. Despite the fact that the program may not be subject to any economies of scale or interregional spillovers, decentralization does not necessarily welfare dominate centralized delivery.

The primary problem with the decentralized delivery mode is the proneness of local governments to pressure from local elites to divert supplies. These reflect weaknesses in the

²³In the centralized case the pattern of grants is decreasing in the local poverty rate, assuming that the coefficient of inequality aversion represented by $\frac{-yv''_p}{v'_p}$ is greater than one.

functioning of a fair electoral process at the local level, lower levels of political awareness among the poor, and the tendency for wealthier groups to form special interest groups that contribute to campaign finance of political parties. The anti-poor bias, and hence targeting failures under decentralization, tend to be more severe in regions with high poverty rates. The centralized mode is vulnerable to similar forms of political capture at the national level, in addition to problems of black market diversion and bureaucratic corruption resulting from weak monitoring and incentive systems. However the cross-sectional pattern of targeting failures gets reversed: high poverty regions are prone to lower intra-community targeting failures under centralization because of the smaller scale of black markets. Hence intra-regional targeting improves with decentralization in low poverty regions, and worsens in high poverty regions.

In our model decentralization enhances the cost-effectiveness of the delivery system, and reduces bureaucratic corruption. Neither system is successful in providing local communities with insurance with respect to random variations in region-wide need or service cost that cannot be verified by the central government. The decentralized system also tends to exhibit superior intra-regional targeting efficiency. However, it may be accompanied by a smaller scale of delivery compared with the centralized model; the cutbacks are likely to be particularly severe in poorer regions, resulting in a poorer inter-regional targeting performance.

It is our hope that these predictions can be tested empirically in future research. Evidence for the connection between targeting and decentralization does not appear to be currently available.²⁴ The existing literature does however provide evidence concerning the relation of targeting of antipoverty programs to the scale of the program and the local poverty rate for specific countries or programs: e.g., Lanjouw and Ravallion (1999) for India, Ravallion (1999) for Argentina, and Galasso and Ravallion (2000) for Bangladesh. These papers generally find that intra-regional targeting performance improves with the scale of the program, which is consistent with the pattern predicted by our model for the centralized regime.²⁵ In the case of the Bangladesh education program, Galasso and Ravallion (2000)

²⁴For a survey of the existing evidence see Mookherjee (2001).

²⁵The Indian data pertains to National Sample Survey data for 1993-94, a period when the Indian anti-

find inferior targeting performance within communities with greater land inequality, which provides support for the ‘local capture’ phenomenon that underlies our model.

For the sake of theoretical clarity, our analysis made a number of simplifying assumptions. These include assumptions concerning only two categories of citizens within each community, absence of trade across different communities, complete non-participation of the poor in election campaign activities, inelastic market demand for the good or service in question, self-interested behavior and absence of disciplinary mechanisms for centrally appointed bureaucrats, absence of self-targeting mechanisms or matching grants that would permit some community insurance, absence of any information at the central government level concerning local needs, costs or service allocations, and so on. Many of these assumptions may not be applicable in any real situation. Our model should be viewed as a convenient benchmark, where we derived the consequences of one set of assumptions. It can hopefully be used to explore the effects of variations in assumptions that seem pertinent to particular applications. We leave such explorations to future research.

poverty programs tended to be centralized by and large. So the findings of Lanjouw and Ravallion are consistent with our model. The Bangladesh program studied by Galasso and Ravallion on the other hand pertains to a decentralized program, but in this case the effect of scale on targeting happens to be statistically insignificant, so does not provide strong evidence against our model. The Argentinian evidence studied in Ravallion (1999) concerns an employment program which appears to have an important decentralized administered component, but the evidence here deals with targeting across rather than within communities, so does not relate to our predictions concerning intra-community targeting.

Acknowledgments

This is a revised version of Part 1 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 and the National Science Foundation Grant No SBS-9709254. For helpful comments we thank two anonymous referees, participants of the MacArthur Research Network on Inequality and Economic Performance, the 1999 Annual World Bank Conference on Development Economics, and the ISPE Public Finance and Development Conference at Cornell University.

References

- Allingham M.G. and Sandmo A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*. Reprinted in *Modern public finance*. Volume 1. International Library of Critical Writings in Economics, no. 15, Aldershot, U.K. and Brookfield, Vt.: Elgar, 1991, pages 50-65.
- Banerjee, A., 1997. A Theory of Misgovernance. *Quarterly Journal of Economics*, 62, 1289-1332.
- Bardhan, P., 1996a. Efficiency, Equity and Poverty Alleviation: Policy Issues in Less Developed Countries. *Economic Journal*, 106, 1344-56.
- Bardhan, P., 1996b. Decentralized Development. *Indian Economic Review*, 31 (2), 139-56.
- Bardhan, P. and Mookherjee, D., 1999. Relative Capture of Local and National Governments: An Essay in the Political Economy of Decentralization. Working Paper, Department of Economics, University of California, Berkeley.
- Bardhan, P. and Mookherjee, D., 2000. Capture and Governance at Local and National Levels. *American Economic Review*, May 2000.

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

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

Conning J. and Kevane M., 1999. Community Based Targeting Mechanisms for Social Safety Nets. Mimeo, Department of Economics, Williams College.

Crook R. and Manor J., 1998. Democracy and Decentralization in South Asia and West Africa. Cambridge: Cambridge University Press.

Dasgupta A., 1997. Decentralisation Experiments in the State of West Bengal in India. Ministry of Finance, Government of West Bengal, Calcutta.

Delli Carpini M. and Keeter S., 1996. What Americans Know About Politics and why it matters. New Haven: Yale University Press.

Gaiha R., Kaushik P.D. and Kulkarni V., 1998. "Jawahar Rozgar Yojana, Panchayats, and the Rural Poor", *Asian Survey*, vol.38.

Gaiha R., Kaushik P.D. and Kulkarni V., 2000. Participation or Empowerment of the Rural Poor: The Case of Panchayats in India. In V. Damodaran and M.Unnithan-Kumar (Edited), *Post Colonial India*, Manohar, Delhi, 2000.

Galasso E. and Ravallion M., 2000. Distributional Outcomes of a Decentralized Welfare Program. Working Paper no. 2316, World Bank Research Department.

Green J. and Laffont J.J., 1986. Incentive Theory with Data Compression. In Heller W., Starr R. and Starrett D. (Edited), *Essays in Honor of K.J. Arrow*, Vol. 3, Cambridge University Press, 1986.

Green J. and Laffont J.J., 1987. Limited Communication and Incentive Compatibility. In Groves T., Radner R. and Reiter S. (Edited) *Information, Incentives and Economic*

Mechanisms, University of Minnesota Press, 1987.

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

Hamilton A., Madison J. and Jay J., 1937. *The Federalist* (originally published, 1787). New York: Tudor.

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

Jimenez E. and Sawada Y., 1999. Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program. *World Bank Economic Review*, September 1999, 13(3), 415-442.

Laffont J.J. and Martimort D., 1998. Collusion and Delegation. *Rand Journal of Economics*, Summer 1998, 280-305.

Lanjouw P. and Ravallion M., 1999. Benefit Incidence and the Timing of Program Capture. *World Bank Economic Review*, 13(2), May 1999, pp.257-274.

Melumad N., Mookherjee D. and Reichelstein S., 1992. A Theory of Responsibility Centers. *Journal of Accounting and Economics*, Vol. 15, pp. 445-489.

Melumad N., Mookherjee D. and Reichelstein S., 1997. Contract Complexity, Incentives and the Value of Delegation. *Journal of Economics and Management Strategy*, vol 6(2), Summer 1997, 257-289.

Mookherjee D., 1998. Incentive Reforms in Developing Country Bureaucracies: Lessons from Tax Administration. In 1997 Annual World Bank Conference in Development Economics, Stiglitz J. and Pleskovic B. (Ed.), World Bank.

Mookherjee D., 2001. Combating the Crisis in Government Accountability: A Review of Recent International Experience. Working Paper, Institute for Economic Development, Boston University.

- Oates W. E., 1972. *Fiscal Federalism*. Harcourt, Brace and Jovanovich, New York.
- Ravallion M., 1999. Is More Targeting Consistent with Less Spending? *International Tax and Public Finance*, 6, pp. 411-419.
- Rose-Ackerman S., 1999. *Corruption and government: Causes, consequences, and reform*. Cambridge; New York and Melbourne: Cambridge University Press.
- Rosenstone S. and Hansen J., 1993. *Mobilization, participation and democracy in America*. New York: Macmillan.
- Seabright P., 1996. Accountability and Decentralization in Government: An Incomplete Contracts Model. *European Economic Review*, vol. 40.
- Tendler J., 1997. *Good government in the tropics*. Johns Hopkins Studies in Development. Baltimore and London: Johns Hopkins University Press.
- Wade R. (1985). *The Market for Public Office: Why the Indian State Is Not Better at Development*. *World Development*. Reprinted in Fiorentini G. and Zamagni S. (eds.) *The economics of corruption and illegal markets*. Volume 1. *The economics of corruption*. International Library of Critical Writings in Economics, Cheltenham, U.K. and Northampton, Mass.: Elgar.
- Wade R., 1997. *How Infrastructure Agencies Motivate Staff: Canal Irrigation in India and the Republic of Korea*. In Ashoka Mody (Ed.), *Infrastructure Strategies in East Asia*. Washington D.C.: World Bank.
- World Bank, 2000. *Entering the 21st Century, World Development Report 1999/2000*. Washington D.C.: World Bank.
- Yitzhaki S. (1974). A Note on Income Tax Evasion: A Theoretical Analysis. *Journal-of-Public-Economics*; 3(2), May 1974, 201-02.

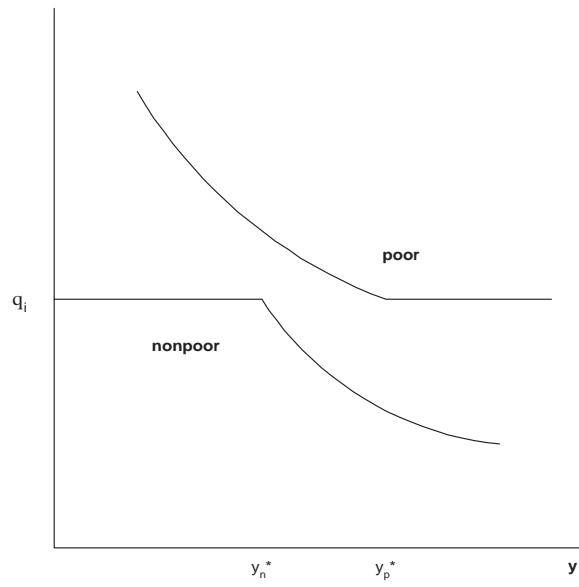


Figure 1: VALUE OF SERVICE TO THE POOR AND NONPOOR

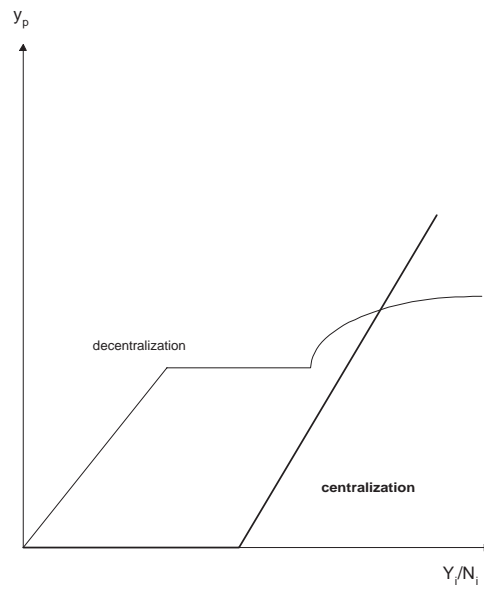


Figure 2: INTRACOMMUNITY TARGETING

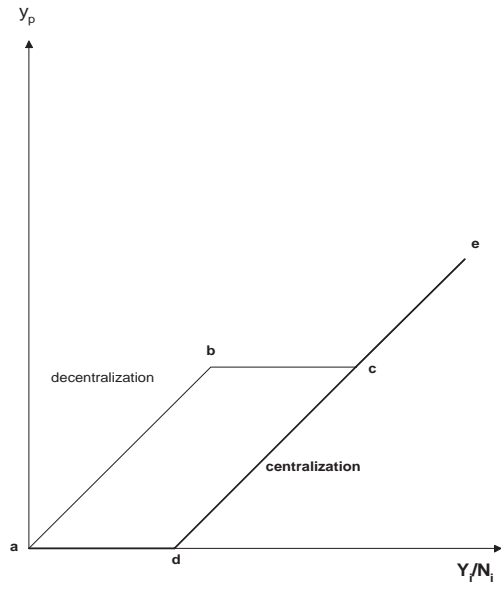


Figure 3: DETERMINISTIC NONPOOR DEMAND

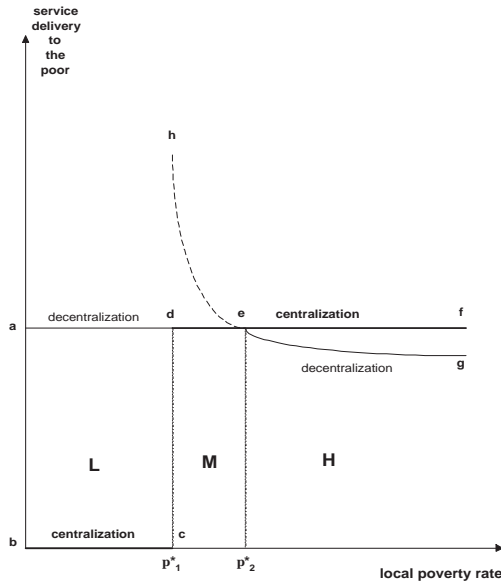


Figure 4: EQUILIBRIUM SERVICE DELIVERY TO THE POOR IN REGIONS WITH VARYING POVERTY RATES

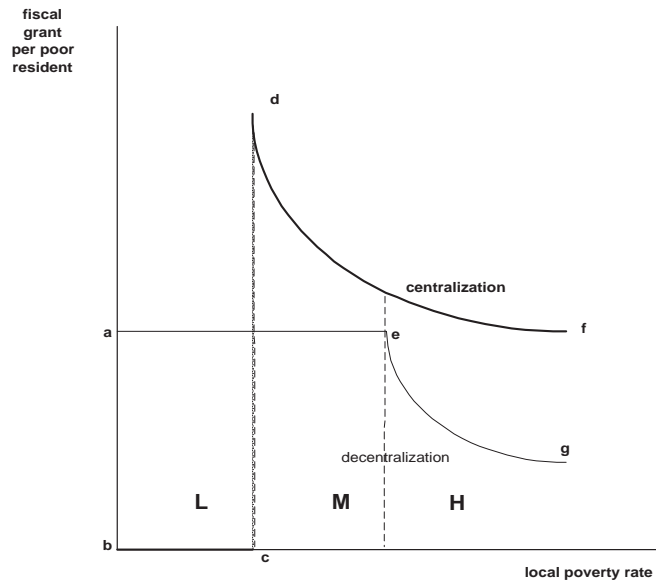


Figure 5: EQUILIBRIUM FISCAL GRANTS IN REGIONS WITH VARYING POVERTY RATES

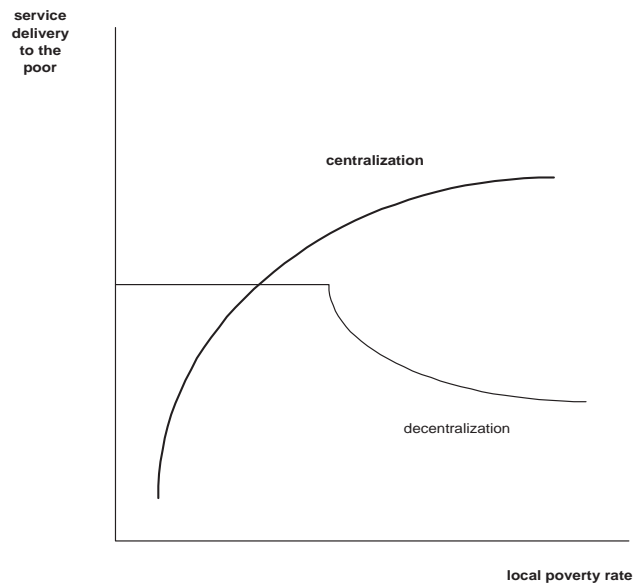


Figure 6: EQUILIBRIUM SERVICE DELIVERY PATTERNS WITH MONITORING OF BUREAUCRATS

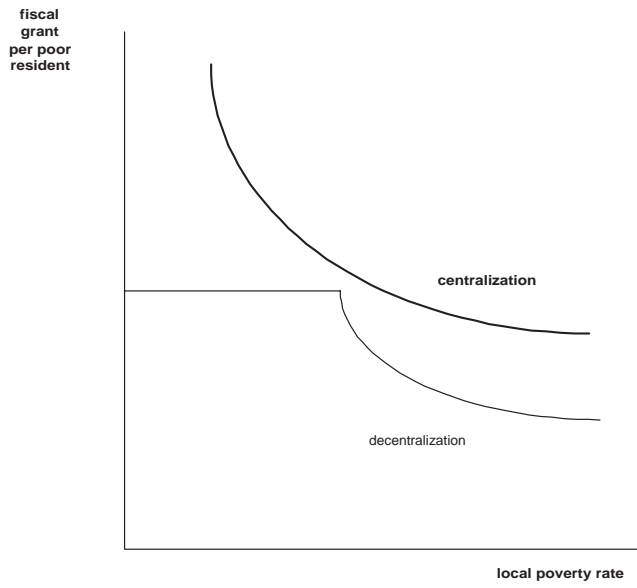


Figure 7: EQUILIBRIUM FISCAL GRANTS WITH MONITORING OF BUREAUCRATS