Caste, Community and Collective Action: The Political Economy of Public Good Provision in India*

Abhijit Banerjee
Department of Economics, MIT

Rohini Somanathan
Department of Economics, University of Michigan

September 20, 2001

*Preliminary and Incomplete.

1 Introduction

One of the abiding puzzles in development economics is the remarkable failure of third world governments to deliver public goods to their people.[1] In 1995, only 13% of the people in Cambodia, 34% in Uganda and 60% in Pakistan had access to safe drinking water.[2] Only 28% of one-year-olds in Chad, 38% percent of one-year-olds in Kenya, and 44% of the one-year-olds in Papua New Guinea were immunized against measles.[3] As late 1991, that is after more than forty years of having an independent and democratically elected government, 27% of villages in India did not have a primary school, 67% did not have any health infrastructure and 31% had no electricity.[4]

Yet there is widespread agreement that these interventions are worthwhile. Micro-studies that estimate the social and private returns to public intervention in health and primary education typically conclude that the net benefits are substantial.[5] There is also little doubt, for most of these public goods, that the potential beneficiaries want them. Bates (1973) claims that the "promise of universal primary education is a political necessity in most African states" and that the absence of primary schools led to demonstrations in many Zambian townships in 1969. In the National Election Study, a post election survey of voters following the parliamentary elections in India in 1996, four out of the top ten major problems facing the country are related to physical or social infrastructure.[6]

While one can think of specific reasons for why individual countries have failed (civil wars, mad dictators), a number of these poorly performing countries have had many years of relatively stable populist, if not democratic, government. Income differences, moreover, explain only a fraction of the enormous variation in the availability and the quality of these goods. We know from the work of Amartya Sen and others that Sri Lanka has public health and primary education statistics comparable many rich countries despite being a low-income country,[7] and within India, Punjab, which ranks among the richest states in the country, has infant mortality rates that are almost three times those of Kerala, a southern state with modest per capita income.[8]

One response to this evidence of widespread failure has been to question the

[1] Public goods, here and in the rest of the paper, refer, not to pure public goods, but to goods that are publicly provided and have decreasing marginal costs over large ranges of provision.

[2] Source: World Development Report, 1998/99.

[3] Source: Human Development Report, 1999.

[4] Census of India, 1991, Village and Town Directory data.

[5] See, for example, Duflo (1999) on the benefits from the Indonesian primary school construaction program in the 1970s or Shultz (2001) on the return to school attendance for poor children in Mexico.

[6] 11.1% of those surveyed report drinking water as a major problem facing the country, 8.5% report education and health, 7.7% report transport and communication and 6.6% report electricity (Mitra and Singh, 1999, pages 107-108).

[7] Dreze and Sen (1995).

[8] Dasgupta (1993) p.314

presumption that popular control or democracy is enough to make the state do what its people want and a corresponding shift to the view that we need to better understand the determinants of successful collective action at the local level. Specifically a lot of attention has been paid to the question of whether the social structure of communities, defined in terms of the size and number of different social groups, has any influence on the provision of publicly provided goods.

There are several potential reasons why social composition may affect the nature of collective action. First, tastes might differ across social groups and, as a result, a heterogeneous community may find it harder to articulate a common demand for a shared good. This could be reflected in lower local contributions for these goods, or political activity which favors rent seeking relative to the cooperation. [9] Second, even if everyone wants the same public good, people may have conflicting views of how the private benefits associated with generation of the public good should be shared—who should be in charge of procurement, whose brother should get the construction contract, etc.[10] Third, privileged groups may fear the social mobility brought about by the greater availability of some of these goods, particularly public schools.[11] It may also be the case that community enforcement mechanisms are weaker in more heterogeneous communities and the agency aspects of collective action are therefore more problematic.[12] Monetary contributions, for example, may not be forthcoming on the grounds that the money could be easily expropriated.[13] Or, the different groups may simply dislike each other and avoid working together.[14]

Since local cooperation is difficult to observe and measure, most studies have estimated a reduced form, looking directly at the relationship between measures of community heterogeneity and public goods. Alesina, Baqir and Easterly (1999), Dayton-Johnson (1999), Lam (1998), Miguel (2000), Khwaja (2001) among others, present evidence showing that the supply of particular public goods is negatively correlated with ethnic heterogeneity in the population.

There are however several quite basic problems with the interpretation of this particular relationship. First, can we be sure that heterogeneous communities are not simply substituting certain other public goods for the ones that they are under-supplying. In fact, prima facie, it is entirely plausible that while heterogeneity undermines some types of alliances, it also makes certain other types alliances of more likely, which might favor specific public goods. Thus, it has often been argued that the weakness of working-class political parties in In-

[9] A formalization of this argument is in Alesina, Baqir and Easterly (1999). Esteban and Ray (1999) contains an alternative preference based model.

[10] Banerjee et al (2001) argue that it is conflict over the sharing of private benefits that limits the effectiveness of sugar cooperatives in Maharastra.

[11] Weiner (1995) argues that social elites in India have resisted compulsory primary education because this would make it harder to maintain "differentiations among social classes".

[12] Guggerty and Miguel (2000) make a version of this argument.

[13] Khwaja(2000) observes that this is an important factor in determining capital contributions for the maintenance of infrastructure in the himalayan villages of East Pakistan. For a related but different example, see Guggerty and Miguel (2000).

[14] Alesina and La Ferrara (2000) make this argument.

dia reflects the fact that caste loyalties are too strong. Increased heterogeneity, by making it less likely that the worker and his boss belong to the same caste, might make class politics easier, thereby promoting the kinds of public goods that workers care about.[15] This kind of substitution would not be a problem for the theory if these substituted goods were clearly less useful(indeed given that there is often a fixed amount one can spend, some substitution may be inevitable), but how do we establish that? To take an example, Alesina, Baqir and Easterly (1999), who were the first to run this kind of regression, actually found evidence of substitution: In their data (which is from the US) ethnic diversity leads to a reduction in spending on sewerage and trash pickup, education, welfare, fire protection and roads, but it leads to an *increase* in spending on health and police and it is not clear that those are less useful.[16]

The other equally basic problem comes from the need to establish the exogeneity of the measures of heterogeneity. A number of the factors that influence heterogeneity (such as urbanization, being in a border area, being near a major road or river, being next to a region where there was a war and therefore a large exodus), also directly influence other economic outcomes including the demand for and supply of public goods. Moreover there is the danger of reverse causality: the poor, for example, may converge to an area which is effective in delivering public goods for the poor, making the area much more homogenous than it would be otherwise. In fact, residential mobility based on differences in public good quality has been a central assumption in the theoretical literature on community composition and the distribution of local public goods.[17] Given that the measures of heterogeneity in these studies are usually more or less contemporaneous with the measures of public good supply (Alesina, Baqir and Easterly (1999) for example, use heterogeneity data from the 1990 census and expenditure data from the 1994 City and County compendium) this is likely to be a serious problem, especially in high mobility environments. Miguel (2000) finds that this type of selection effects can be strong enough to reverse the sign in a regression of the quality of public schools on the heterogeneity of the children who go to that school.

The authors in this literature recognize these problems but given the data limitations, the progress along this dimension has been rather limited. Alesina, Baqir and Easterly try to address the endogeneity issue by using community fixed effects, but once they include fixed effects as well as all their controls, the effect of heterogeneity becomes insignificant or even positive.

In this paper, we a data set from India examine the supply of publicly provided goods in rural areas in the 1970s and 1980s. The data set, put together from various Indian censuses, gives us the fraction of villages in over 300 districts

[15] Of course, it could also go the other way: Workers may find it impossible to work with people from other castes, in which case class politics is only possibl when all workers are from the ssam caste.

[16] The fact other studies do not find similar evidence of substitution might reflect the fact that they only look at a subset of public goods, and it is not implausible, given that they can only look at some public goods, that they would focus on the ones where they have reason to believe that they will find what they are looking for.

[17] For examples, see Benabou (1993), Nechyba (), Epple and Romano ().

in 17 states of India that, in the 1970s and 1980s had a particular public good, as well as measures of the sizes of different caste and religious groups in the district. While our basic empirical strategy follows the literature in regressing the supply of public goods at the district level on measures of social heterogeneity (once again at the district level), specific features of our data allow us to make some progress on the both the substitution issue and the identification problem. As far as the first problem is concerned, we have the advantage of having data on essentially every publicly provided good. This allows us to identify the exact pattern of substitution. We can then ask whether there is evidence that more heterogeneous communities substitute towards less "valuable" public goods. To check on this we use the fact that within the Indian context we can identify a socially group that is known to be politically weak and therefore likely to have limited access to the most valuable public goods. This is the community that in India is called the scheduled tribes (ST). Assuming that we are right about the STs being weak, districts with a higher share of STs should have less of the more desirable public goods. Our theory then predicts that the heterogeneity should have an negative effect on the supply of the desirable public goods identified in this way, and a zero or positive on the supply of other public goods. While this approach has its limitations, we feel that it represents progress on an important issue. On the identification problem, we feel that the nature of data allows to be relatively confident about the reverse causality issue. Our heterogeneity data is based on the 1931 census of India (because, for one, later data on caste is not available). All the historical evidence on migration in India suggests that there extremely little migration out of rural areas in the pre-1931 period,[18] and in any case, much of the expansion of public goods in rural India took place after 1931. We are much less confident that we have resolved the omitted variable problem because while we can control for a district characteristics (rainfall, population density, land distribution, bank deposits per capita, state effects), we cannot, for example, include district fixed effects or instrument for heterogeneity. What we can do is to proxy a district effect by the supply of the one private good that is available from the census—the fraction of villages in the district that have a registered medical practitioner. If the main source of bias comes from a single district effect that alters the supply of every public good as well as this particular private good (say, the level of development, or the wealth of the district), then including a measure of the supply of the private good in the public good regression would reduce or even eliminate the bias. It therefore provides a useful check on our results.

Our results broadly support the view that heterogeneity undermines the ability to get public goods. The 1981 census gives data on 33 public goods and 1991 census gives data for 42 public goods. Out of these 75 goods the coefficient on heterogeneity is negative and significant in about 31, while it is positive and significant in 9 (in the rest it is neither). Moreover, the goods for which the coefficient on the share of scheduled tribes in population is negative and significant are much more likely to also have a negative and significant

[18] See Weiner ().

coefficient on heterogeneity than goods where the coefficient on the share of scheduled tribes is positive. This and other evidence supports the view that the more heterogeneous communities tend to be politically relatively weak (like the scheduled tribes) and therefore less likely to get the goods they want (and conversely more likely to get some of the inferior substitutes). To try to ensure that the results do not come from some omitted variable, we include state dummies in all the regressions. We also control for rainfall and population density, two district characteristics which obviously affect the priority given to different types of public goods. We experiment with using bank deposits per capita as an additional control, to try to control for average wealth. As a further check we control for access to registered medical practitioners (as already mentioned, a private good): This does not change our results.

The credibility of this kind of result depends heavily on whether there are actual (as against potential) mechanisms that plausibly link heterogeneity to public good delivery. We therefore ask whether heterogeneity has a measurable effect on the political process. Using data from state legislative assembly elections for the seventies, eighties and early nineties, we find that districts that are more heterogeneous according to our measure, are also more politically fragmented: Elections in these districts are characterized by a larger number of contestants and a smaller vote share for the winning party.

The final step in making this case is to check whether heterogeneity affects the quality of life in these districts. In other words, it is possible that even though heterogeneity clearly affects the supply of public goods in these districts, it does not really affect the quality of life, say, because these public goods are worthless. We therefore estimate the effects of fragmentation on literacy rates and crime rates and find, in both cases, negative effects of increased heterogeneity.

The next section provides some background information on the institutions which influence government expenditures and briefly examines the growth of infrastructural facilities over the period of our study. Section 3 describes our data. Section 4 outlines our empirical strategy and presents our results. Section 5 reports results on political outcomes which suggest that at least a part of the effect of heterogeneity operates through its impact on political behavior. In conclusion we briefly look at the effect of heterogeneity on several measures of the quality of life including literacy, infant mortality and crime.

2 The Institutional Setting

The Indian constitution divides government functions and financial authority between the central and state governments. As in many federal systems, it is the states that are primarily responsible for health, education and various kinds of community development programs. State expenditures form about 80 per cent of the total government expenditure in these categories.[19]. The states receive large financial transfers from the center, to allow them to spend

[19] Govinda Rao and Nirvikar Singh (1998).

much more than they can tax. A significant part of these transfers are made to implement development programs that are outlined in the national five year plans. Local governments, until recently, have been created at the discretion of individual states and though many in number, they have had little control over taxes or expenditures over the period of our study.[20] Turning to trends in the growth of infrastructural facilities, we see an expansion of primary and secondary education in the fifties and sixties, but little investment in other infrastructure in rural areas. The five year plans formulated during this period emphasized the growth of heavy industry through large investments in state-owned industry. As a result, in 1971, while 52 per cent of all Indian villages had primary schools, only 25 per cent had paved roads, 18 per cent had electricity, 2.5 per cent had tapped water and 6 percent had any medical facilities. In response to these signs of neglect and studies which pointed to high and unchanged levels of rural poverty during sixties, the Minimum Needs Programme was introduced as part of the fifth five year plan in 1974.[21] It sought to bring an elementary school within a one-mile radius of every child and to greatly expand roads, electricity, water and health facilities in villages.

There was indeed a dramatic expansion in many of these facilities over the next two decades, but at least as striking as the overall expansion was the unevenness with which these facilities were distributed both across and within states. The states of Bihar, Gujarat and Haryana all had about 15% of their villages with electricity in 1971. By 1991, almost every village in the Gujarat and Haryana had electricity, while the figure for Bihar had only risen to 40%. Disparities across districts within most states in 1991 were also large. Almost every village in the district of Guntur in Andhra Pradesh had a primary school, electricity and a paved road, while less than half the villages in Vishakapatnam district of the same state had these facilities.[22] Table 0 provides some relevant descriptive statistics: It is notable that even in 1991, the range of almost every public good variable starts at zero and streches to over 90%. There were entire districts which did not have a single middle school, and districts where not a single village had tapped water and districts where not a sngle village had a paved road. Moreover the mean level for these three variables were, respectively, .26, .28 and .45 sand the maximum was 100% in all three cases. Literacy rates in 1991 varied between 3% and 83% for women and 165 and 86% for men and the number of murders went from 0 to 267.

A relevant feature of structure of government spending in India is that locally provided goods are not locally financed. While the process by which governments allocate these goods is not transparent and probably quite complex, it is plausible that the ability of communities to collectively articulate their demands to politicians and administrators is important in determining their location. Such articulation may involve informing politicians that the provision

[20] A constitutional ammendment in 1993 forced all states to form village level governments which would be elected every five years. Since then their share of government spending has increased to about 6%. Their fiscal powers are still very limited.

[21] Chaudhuri (1979) surveys estimates of rural poverty during the sixties.

[22] The state of Kerela is the one exception here, with uniformly high levels of provision.

of these goods is important for their reelection, visits to district administrators who implement development plans or it may mean local contributions of land or labor which allow the building of schools and health facilities. The south Indian village studied by Epstein et al.(1998), had many facilities that were unusual for a village of its size-they obtained piped water as a result of meetings between the village council and the district administration, a high school was sanctioned after one of village families provided a building, and a health centre was constructed after the villagers donated land for it. The collective effort involved in all these activities may well be influenced by the heterogeneity of local communities.

How much heterogeneity is there and why might it matter? The social structure of Indian villages and its effects on village life has been intensely studied by anthropologists, to the extent that the Indian caste system has functioned as the primary lens through which village life has been observed. Hindus (the major religious group) are divided into a number of castes, with strict and long-standing rules which govern their interaction. Marriages rarely take place across caste boundaries and the sharing of food and other social interaction dictated by the caste system. While there is some slow mobility of caste groups in the hierarchy over long periods of time, there is almost no mobility of individuals across these groups.[23] Within villages, castes often inhabit different hamlets and the distinction between upper and lower castes is particularly sharp. Muslims and Christians form the major minority religious groups and there are similar rules governing contact with them. In terms of magnitudes, our measure of heterogeneity (the standard ethnic fractionalization index) has a mean around .85, compared to the mean value of .26 for US cities that Alesina, Baqir and Easterly report and moreover it stretches from around 0.03 to 0.998. The share of scheduled tribes goes from 0 to 98.5% in 1991 and that of scheduled castes from 0 to 54.5%. The gini coeffcient of the land distribution goes from a very low 0.127 to a impressively high 0.85.

In this context of sharp divides between different groups, many of the effects of heterogeneity alluded to in the last section may be present. There is anecdotal evidence, mainly from village studies, that different castes often use different water sources and, because of norms that limit the entry of lower castes into upper caste neighborhoods, the location of schools and other public places within villages influences their use. Changes in village leadership have often been accompanied by a change in the location of the public building where village meetings are held.[24] Political parties often pledge allegiance to partic-

[23] "Classes are- in principle and, to some extent, in practice- open; castes are not. One may change one's position from tenant to landowner, or from agricultural labourer to owner-cultivator. One cannot, however, change from a Vellala into a Brahmin or from a Palla into a Vellala...Movement upwards or downwards within the caste system is, in theory, inadmissible, although there is some movement in practice..Yet there are significant differences between social mobility in the caste system and social mobility in the class system. In the latter it is the individual who moves up or down, whereas in the former entire communities change their position" Beteille, p. 190.

[24] Beteille (1969) in his classic study of a village in south India describes the way in which temples, meeting rooms and the elementary school are strategically located to make access

ular caste groups and their interests and in fact many have argued increased rural involvement in politics has sharpened social caste and communal divides.[25] Fukunaga(1993), in his study of village factions in Uttar Pradesh, describes how development projects introduced in the 1970s created large potential rents for those who administered them in the village. This led to increased political activity which was almost always along group lines. The lower castes, who had shown little interest in earlier elections were now mobilized by those seeking political office.

Our objective here is to investigate whether these descriptions capture processes that were at work more broadly across the country and to quantify their relative effects on the provision of public goods.

3 Data and Empirical Strategy

3.1 A Description of Data Sources

Our data on the location of publicly provided goods comes from the village directories of the Census of India. These list a variety of village amenities, including schools, medical facilities other forms of physical infrastructure. We use data from the 2 census years, 1981 and 1991. In our estimation we use aggregate facilities at the district level since we have no information on social structure for individual villages. We leave out three of the public goods listed in the census—rivers, lakes and springs–on the grounds that these not publicly provided. In addition to listing most public amenities, the census also lists the number of villages in each district with private medical practitioners. We use this in our identification strategy (described below) to control for district fixed effects. We use data on population shares of caste and religious groups to construct measures of social fragmentation. Self reported data on religion is available for each census year, but detailed caste data is not available for the post-independence period. Census reports in this period contain only the number of *scheduled castes* and *scheduled tribes.* These are groups that have been considered particularly under privileged as a result of their position in the social hierarchy and have benefited from various forms of affirmative action since independence. Data is available for 397 districts in 1981 and 446 in 1991. The different numbers of districts in different years results in part from the creation of new districts but also because the census was not held in certain states in certain years due to political unrest.[26]

easy for particular castes. He also finds that social clevages are heightened as villages partipate in political acitivity. "As the election campaign mounts, people tend to identify themselves progressively with one part or another. The cleavages within the village community are more sharply focussed, and the links between political interest and social structure are brought to the surface" (p. 179)

[25] "In Rampura in 1948, inter-caste relations were on the whole cooperative if not friendly...But with the introduction of adult franchise and of the electoral principle into panchayats and other local self-governing institutions, tensions between the castes increased sharply" (Srinivas, p150). Also see Singh (1993).

[26] Assam in 1981 and Jammu and Kashmir in 1991.

The last enumeration of more detailed data on caste that we have is from the 1931 census, and it is a modified version of these data that we combine with current data on religions, to construct our measures of heterogeneity.[27] The data is available by districts, separately for each of the British Indian provinces and princely states. While state boundaries were redrawn after independence, district boundaries remained more or less intact and we can therefore use this data to construct caste shares for current districts. For new districts created by subdividing old ones, we weight the caste figures from the original district according to the area of the new district which was taken from them.

The number of castes listed in the 1931 is very large and we restrict ourselves to Hindu castes which form more than 1% of the population of each state or province in 1931. Putting data for different states together, we have a total of 185 caste groups. We make one major adjustment to this data to account for the increase in the proportion of Hindus after 1931. Some districts had significant Muslim populations which emigrated to the newly created nation of Pakistan around the time of Indian independence in 1947. We scale up the numbers in each caste group, based on the population share of hindus in the current census. This assumes that within Hindus, different castes grew at similar rates over time.[28] To measure of heterogeneity, we mainly use the fractionalization index,

$$h = 1 - \sum_{1}^{n} s_i^2 \tag{1}$$

where s_i refers to the population share of the ith group.

We use data from a variety of other sources to control for other characteristics of districts which affect public good provision. We construct measures of land inequality from data on the number and size of operational holdings provided by the agricultural census of India. Data on rainfall and the bank deposits per capita and other control variables is from the district profiles put together by the Centre for Monitoring the Indian Economy.

The electoral data we use is from the election commission and is by assembly constituencies. There are a total of about 3000 assembly constituencies in India, and therefore several in each district. We have matched the boundaries of assembly constituencies to districts in order to generate district level electoral data. State assembly elections are not, in general held in census years and also vary across states. We use the 3 election years closest to our census years. For most states, these are 1972, 1980 and 1991. We use two principal kinds of electoral data- voter turnouts and the shares of total votes received by individual parties. To obtain turnouts, we sum the total votes and eligible voters for all the constituencies of the district; for constructing variables relating to the shares of individual parties, we use average shares across the constituencies of a district.

[27] Some caste data was collected by the 1941 census but it was never tabulated- a combined effect of the war combined with the volatile political situation in India.

[28] While this is certainly not true of urban India, we hope that this is a reasonable approximation for rural areas. There is very little evidence on group wise variation in fertility and migration rates for us to be able to do much else.

As a measure of political fragmentation, we use a factionalization index similar to the heterogeneity index above, replacing population shares of different groups by the share of total votes received by individual parties. We average across constituencies in a district to get the index of political fragmentation for the district. Data on literacy are from the Census and crime data are based on police records published by the Central Crime Records Bureau.

3.2 Empirical Strategy

The basic cross sectional relationship estimated in this paper takes the form

$$y_{ijt} = f_i(h_{jt}, \mathbf{x}_{jt}, \varepsilon_{ijt}) \tag{2}$$

where y_{ijt} is the extent of provision of the ith public good in the jth district in census year t, h_{jt} is a measure of heterogeneity in the district, $\mathbf{x}_{jt}$ is a vector of other district characteristics and ε_{ijt} is a district and good specific shock. The problems in interpreting the results of estimating this relationship come from three main sources—substitution, reverse causation and omitted variables. We discuss them in order.

3.2.1 Substitution

The theory implicit in our discussion predicts that more heterogeneous districts will have less access to the public goods that they want. This does not automatically imply that they will get less of every public good. In fact they may be given more of some public goods precisely because they have less of others. This may be, for example, a part of some political mechanism aimed at making sure that they do not become too unhappy. Or it could be a part of an attempt to deal with the consequences of the lack of other public goods. For example, a village that has been traditionally under supplied with educational infrastructure might need and get an adult education center ahead of a more favored district.

It follows that the testable prediction of the theory ought not be that the coefficient on heterogeneity is negative on every public good. On the other hand, it is obviously not easy to test a hypothesis that predicts that a sign will be sometimes positive and sometimes negative. We therefore need some way of identifying those goods where we would expect a negative coefficient. These ought to be the goods that are most desirable but, once again, we do not always have a priori information on how people rank these goods.

One admittedly crude way to get around this problem is to use the fact that we know that the scheduled tribes in India tend to be the single weakest political group of any size and therefore are unlikely to have much access to the most desirable public goods.[29] This tells to expect to see a negative sign

[29] This ought to be true at least once we control for the share of christians in the district, since scheduled tribes have a large fraction of christians and therefore have acces to the kinds of public goods that christian missionaries bring with them.

on heterogeneity whenever the coefficient on the population share of scheduled tribes is negative but not necessarily otherwise. Of course, we are unlikely to find such an exact relation in the data, given that public goods vary along other dimensions as well, but it is useful to check whether the predicted pattern broadly accords with the data.

3.2.2 Reverse Causation

Reverse causation can arise through at least two different routes. First, as suggested above, the successful provision of a certain type of public good will attract those who are heavy users of that particular public good. Second, ethnic identities are partially formed in the process of political mobilization—people often embrace specific ethnic identities when ethnicity becomes politically relevant. Therefore a higher degree of ethnic heterogeneity may reflect more political mobilization along ethnic lines, which, for example, may be a result of the fact that the area lacks public goods.

The fact that our measure of heterogeneity is based on data from the 1920s makes us relatively less vulnerable to this particular problem. Most of the expansion in the supply of goods took place after independence and also India until the 1970s was known as a very low mobility economy and this was even more true in the 1920s. It is therefore unlikely that population composition was substantially altered by the presence of public goods. There was also very little political mobilization along ethnic lines in India before the 1930s— caste-based movements such as the Dravidian movement, only gathered force in the 1930s.[30] The one possible exception was the mobilization of Muslims during the Khilafat movement, though even there the focus was on influencing the national (British) government rather than local governments. However in order to make sure that this is not a problem, we check the robustness of our results using measures of heterogeneity within the non-muslim population.

3.2.3 Omitted Variables

It is not hard to think of omitted variables that will influence the supply of public goods and are correlated with heterogeneity. One possibility is simply that different ethnic groups want different public goods (Brahmins want education, traders want roads etc.) and that the heterogeneity measure is picking up changes in the share of these different groups. To check for this possibility we correlate our measure of ethnic fragmentation with the shares of specific groups. Furthermore, in all our regressions we control for the share of all the major ethnic groups and have tried out variants where we also put in the square of the share of some specific groups. A more plausible variant of this argument takes as its premise that people in areas which are ethnically diverse tend to be different (even after controlling for the shares of each group) and this difference is reflected in their particular tastes for public goods. In other words, ethnically

[30] This is not surprising since the dominant force in politics in this period was the nationalist movement and elected local governments had little or no power till the XX reforms of 1935.

diverse populations can be different from more homogeneous populations even when all ethnic groups have exactly the same preferences on average and there is no other basis for ethnic conflict. Thus an area that is a center of trade will tend to have a population weighted towards traders from everywhere and therefore heterogeneous. It is also likely to have a population that cares relatively more about roads and other means of communication and perhaps less about health centers and water supply. This is an example of taste-based substitution which is clearly different from the kind of substitution that we discussed above. It suggests that the effect of heterogeneity may vary significantly even within the set of desirable public goods. However if this is all that is going on, we would expect a rough balance between the set of desirable goods where heterogeneity is an advantage and those where it is a disadvantage. Another class of potential omitted variables are determinants of the overall political clout of the district such as its wealth, education and urbanization. Districts which score low on this variable may very well end up with less of every desirable public good, in contrast with the previous case where districts only differ in their taste for pubic goods. A possible example of such an omitted variable may be the historical patterns of landownership which may explain both why certain districts are more caste divided than others (caste, after all, grew out of division of labor) and also why they are poorer and therefore less able to get public goods from the state. There is nothing definitive that we can do to rule out this kind of identification problem. We do however control for a state fixed effect, to deal with the fact that different states have very different fiscal resources and also different patterns of ethnic diversity. We also control for rainfall and inequality in the land distribution, which should pick up some of the effects of historical differences in landownership. Finally, we estimate results with and without controlling for literacy levels and population density (we do it both ways because we are concerned about potential endogeneity isuues) and restrict the analysis to rural districts to minimize the effects of urbanization.

Our data also allows us a possible test of the 'favored district' view. As mentioned above, we have data on the number of registered private medical practitioners in the district. Suppose there is one district characteristic, say wealth or a the level of education in the population, that increases the demand for both private and publicly provided goods. Then, under the assumption that the supply of private goods is unaffected by heterogeneity, we show below that it is possible to control for unobserved the district characteristic by including the supply of the private good as an additional control. In the more general case where there are several different unobserved characteristics that affect private and public goods differentially, this control will be imperfect but it should still mitigate the bias.

This particular strategy relies crucially on the assumption that heterogeneity *per se* does not affect the supply of registered private medical practitioners. This seems reasonable since we control for the fact that different groups may have different preferences. We can also look at the coefficient of heterogeneity in the equation for registered medical practitioners. However since we have introduced this variable into the analysis precisely because we feel that the cross-section

regressions may be biased, this, at best, provides limited comfort. One reason why heterogeneity might have an effect on the supply of private medical practitioners is that heterogeneity affects the supply of public medical infrastructure (hospitals, dispensaries, health centers, etc.). Depending on how these function, private registered medical practitioners may be either complements or substitutes for these public goods. In either case this would induce a correlation between the supply of private medical practitioners and heterogeneity. To control for this possibility we also estimate a specification where the supply of the various public medical facilities is included as an additional control.

We also use essentially the same "reduced-form" specification to investigate the effect on the political outcomes and the measures of the quality of life. While a more structural approach (looking at how the effect on political behavior translates into an effect on the supply of public goods, for example) has the potential to be more insightful, the identification issues seem largely intractable.

3.3 Empirical Specification

Our very 'reduced form' model starts from the level of a village since the outcome we are interested in is the share of villages with a particular public good. We assume that each village has a certain amount of political capital which it can use to lay claims on various public goods. The total amount of political capital associated with village k in district j is given by P_{kj}.

The rule for allocating public goods is simple: village k in district j will get public good i if and only if

$$P_{ikj} \geq C_i.$$

We think of village level political capital $P_{ijk} = P_{ij} + \xi_{ikj}$, where P_{ij} is the mean level of political capital in the district for getting good i and ξ_{ikj} is a random village specific term with $E_k[\xi_{ikj}] = 0$ for every (i, j) combination. The probability that village k will have good i can therefore be written as $\Pr\{\xi_{ikj} \geq C_i - P_{ij}\}$.

We now postulate that the average amount of political capital in a district is a function of ethnic heterogeneity (h_j) and other district characteristics ($\mathbf{x}_j$): Specifically let $P_{ij} = \beta_i h_j + \boldsymbol{\gamma}\mathbf{x}_j$, with the presumption that β_i is negative. Now the probability that village k has public good i takes the form

$$\Pr\{\xi_{ikj} \geq C_i - \beta_i h_j - \boldsymbol{\gamma}\mathbf{x}_j\},$$

We assume that this error term follows an extreme value distribution and, for each good separately, we estimate a logit model with grouped data, since we are using district level aggregates. Denoting the proportion of villages with good i by y_{ij}, and the number of villages in district j by n_j, this is equivalent to a linear regression of the log odds ratio for each good, on our explanatory variables. We therefore estimate

$$\log(\frac{y_i}{1-y_i}) = \alpha_i + \beta_i h_j + \boldsymbol{\gamma}_i \mathbf{x}_j + \varepsilon_{ij} \tag{3}$$

separately for 1981 and 1991.

This specification runs into the well-known problem with aggregating logit models unless we can be sure that all villages in the district are identical: In other words, if village characteristics vary within a district then the logit model applies either at the village level or at the district level but not both. An alternative formulation that avoids the aggregation issue but has other obvious problems is the linear probability model:

$$y_i = \alpha_i + \beta_i h_j + \boldsymbol{\gamma}_i \mathbf{x}_j + \varepsilon_{ij} \tag{4}$$

Below, we present estimates using both the logit and the linear probability model. They are qualitatively very similar.

4 Regression Results

4.1 Public Goods

Tables 1A, 1B and 1C present our basic results for 1981 and 1991. We restrict ourselves to the 16 largest states which contained over 98 per cent of the population at the time of the 1991 census.[31] These states contain 322 districts in 1971, 337 in 1981 and 392 in 1991. We begin, in columns 1 and 2, with a very parsimonious specification: The right hand side variable of interest is obviously our measure of social fragmentation. In addition we include state dummies to capture differences across states in the mean level of these goods that are not observable in our data and rainfall to proxy for agricultural productivity. We also control the population shares of brahmins, the three major religious groups (muslims, Christians and Sikhs) and scheduled castes and tribes.

We clearly some support for our hypothesis: Heterogeneity is negative and significant (at the 10% level) for 15 out of the 33 public goods in 1981 and 12 out of the 42 public goods in 1991. Heterogeneity has a positive and significant effect on only 2 1981 goods and 7 1991 goods. We next reestimate the model, including as additional explanatory variables a measure of land inequality in the district, the level of per capita bank deposits and average village population. These results are reported in Columns 3 and 4 in the same table. The coefficients on fragmentation from the two specifications are quite consistent. In the 1981 sample we find a negative and significant coefficient for 16 of the 33 goods, while in the 1991 sample it is for15 out of the 42 goods. A positive and significant coefficient occurs twice in the 1981 sample and 7 times in the 1991 sample

The fact that the negative heterogeneity effect is less strong in the 1991 sample is not unexpected. The 1980s was a period in India when people were

[31] There were a total of 25 states at the time of the 1991 cenus. The 16 largest are, in alphabetical order, Andhra Pradesh, Assam, Bihar, Gujarat, Haryana, Himachal Pradesh, Karnataka, Kerela, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. We exclude the small north eastern border states of Arunachal, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura and the recently creates state of Goa in the West.

becoming more vocal in their demand for public goods and it seems plausible that this would lead to the supply of public goods becoming less unequal. Moreover we see the introduction of a large number of new public goods in this period, including adult literacy centers, health centers, TB clinics, child welfare centers, nursing homes, maternity homes, nalas etc. Only one of these new public goods (maternity home) shows a significantly negative heterogeneity effect, while a number show a positive and significant effect, which is perhaps not surprising given that a number of these goods seem to be intended to compensate for the lack of other public goods (adult literacy centers for primary schools, TB clinics and child welfare centers for primary health centers, nursing homes for hospitals, etc.).

More generally, as already discussed, we can try to distinguish between the goods that everyone wants and the ones that are offered as substitutes using the sign of the ST effect. It is easily checked that when the ST effect is negative, which happens 17 times in both the 1981 sample and the 1991 sample, the heterogeneity effect is negative, respectively, 9 times and 11 times. By contrast when the ST effect is positive, which occurs once in the 1981 sample and 4 times in the 1991 sample, the heterogeneity effect is only significant in one case (in 1991) and then it is positive.

The fact that the ST effect is systematically negative is also important in itself. Moreover the ST effect is often quite large: For example, for tubewell/handpump in 1981 the ST coefficient is -3.36, implying that 1 percentage point increase in the share of ST will reduce the percentage of villages that have this particular public good (mean 16%) by 0.46 percentage points. Or for middle schools in 1991 the coefficient is -0.87 implying that a 1 percentage point increase in the share of STs will reduce the fraction of viallges with middle schools (mean 21%) by 0.11 percentage points. It is clear that the STs are still substantially behind in terms of access to public goods. Interestingly, this appears not to be true of the scheduled castes—the effect of more scheduled castes in the district seem s to be different for different goods but about equally likely to be positive and negative. This is consistent with the increased political mobilization of scheduled castes in independent India and the relative political marginalization of scheduled tribes.

Finally, the estimated effect of inequality is of some independent interest. . The unequal access to resources implied by caste system has meant that different castes experience very different level of material well being (Deshpande 2001). In fact, caste based affirmative action policies have been introduced to address these. It is therefore possible that our caste fragmentation measure is simply a proxy for inequality, which is the relevant social divide. This does not seem to be the case based on our analysis of the data. The inclusion of inequality measures strengthens rather than erodes the effects of social fragmentation- caste and class have separate effects on collective action. The sample correlation between our measures of inequality and social fragmentation is no more than .2 in any year.

It is also worth mentioning in regard to the effects of inequality is that the effects of inequality are sensitive to the particular measure of inequality being

used. While different measures have little effect on the fragmentation coefficient, they have different independent effects on public good provision. As can be seen from Table 1, the gini coefficient is mostly insignificant, but when significant (for example in the case of domestic power supply in 1981, telephone connections in both years and roads, hospitals and medical practitioners in 1991) it is positive. If instead we use measures of inequality that are more sensitive to the lower tail of the income distribution, such as the median to mean land ratio, we find negative coefficients. Taken together these results seem to suggest that the relatively rich can be effective in getting goods to their villages (especially goods like phone connections, which are demanded disproportionately by them), but at the same time, having large numbers of the very poor hurts access to public goods.

Turning to the coefficients on our other explanatory variables, districts with larger villages have more of all goods. And wealthier districts (measured by bank deposits per capita), not unexpectedly, have more of most goods.

4.2 Robustness Checks

Columns 1 and 2 of Tables 2A and 2B report results on the coefficient on heterogeneity when we estimate the linear probability model given by 4 instead of a logit. As is apparent, the results are quite similar.

The effect of our measure of social fragmentation on the availability of private medical practitioners was reported in Table 1. .Since this is a private good, we did not count these results in the above discussion. As one might expect of a private good, in all the specifications and both the years, heterogeneity has no effect on its supply but the fraction of STs has a significant negative effect.

If we take additional step of assuming that heterogeneity truly has no effect on the supply of registered medical practitioners (this is supported by the above evidence, but is a not a necessary consequence, since our regressions may be biased), we can use the information on the supply of private medical practitioners to pick up unobserved district characteristics. In particular suppose the true model is

$$\log(\frac{y_i}{1-y_i}) = \alpha_i + \beta_i h_j + \boldsymbol{\gamma}_i \mathbf{x}_j + \delta_i v_j + \varepsilon_{ij}. \tag{5}$$

for all public goods and

$$\log(\frac{y_{rp}}{1-y_{rp}}) = \alpha_{rp} + \boldsymbol{\gamma}_{rp} \mathbf{x}_j + \delta_{rp} v_j + \varepsilon_{rpj} \tag{6}$$

for the private good (since collective action is not a determinant of the location of the private good).[32] Here ν_j is an unobserved district characteristics which

[32] When specifying the district fixed effect in this way, the question of how this is related to village fixed effects comes up again. In the linear model, we could have different village fixed effects and the district effect could just be the average of these. With the logit specification, village fixed effects would have to be constrained.

is potentially correlated with h_j and therefore a potential source of bias. 6 can be rewritten in the form

$$\nu_j = \frac{1}{\delta_{rp}}[\log(\frac{y_{rp}}{1-y_{rp}}) - \alpha_{rp} - \boldsymbol{\gamma}_{rp}\mathbf{x}_j - \varepsilon_{rpj}] \tag{7}$$

which when substituted in 5 gives us:

$$\log(\frac{y_i}{1-y_i}) = (\alpha_i - \alpha_{rp}/\delta_{rp}) + \beta_i h_j + (\boldsymbol{\gamma}_i - \boldsymbol{\gamma}_{rp}/\delta_{rp})\mathbf{x}_j + \frac{\delta_i}{\delta_{rp}}\log(\frac{y_{rp}}{1-y_{rp}}) + (\varepsilon_{ij} - \varepsilon_{rpj}/\delta_{rp}) \tag{8}$$

Estimating this equation yields consistent estimates of β_i even when h_j and ν_j are correlated. However this method relies on their being a single omitted variable. If there are more than one, this procedure can at best limit the bias.

Columns 3 and 4 of Tables 2A and 2B provides estimates of the heterogeneity coefficient estimated from equation 8 The size of the effects are, in general much larger.

We have already suggested that the fact that the presence of medical practitioners may be correlated with other elements of the health infrastructure might induce a spurious correlation between it and heterogeneity. To control for this possibility, we also run the above regression using the share of villages with public health centers and sub-centers as additional controls. The results, reported in column 3 of Table 2, are essentially unchanged.

5 The magnitude of the effect

The estimated coefficients suggest effects of fragmentation that can be quite large. Table 3A and 3B list these effects for all goods for which the coefficient of fragmentation was significant in either 1981 or1991. These effects have been evaluated at the sample mean for each public good. In 1991, a one-standard deviation increase in our fragmentation measure is associated with a 1 percentage point fall in the percentage of villages with hospitals, a 3 percentage point fall in bus services, a 4 percentage point fall in middle schools, post and telegraph facilities and domestic power, and an 11 percentage point fall in tapped water. To put this in perspective, averaging over all districts, only about 3 per cent ofvillages had hospitals, a quarter had tapped water and middle schools,about 40 per cent had bus services and three-quarters had domestic power in1991.

Another way of looking at these coefficients is to ask how much of the average differences between states is explained by heterogeneity. To do this, we first estimate our logit model and then using the estimated coefficients and the observed values of all other explanatory variables, we simply replace the value of our fragmentation variable by that of Punjab-the state with the lowest average fragmentation (.42). We then look at differences between actual values and those predicted using the above procedure for each district and then construct averages for each state. If all districts in India had the same fragmentation

index of .42, Andhra Pradesh would have a 22 percentage point increase in middle schools, a 5 percentage points increase in high schools, a 0.6 percentage points increase in hospitals and a 7 percentage increase in bus services. This would almost double the number of villages with middle and high schools and increase the number with hospitals by more than a quarter.

6 The effect on political behavior

Differences in the nature of politicians elected are a possible mechanism through which social structure could influence the availability of public goods. If preferences across groups for public goods converge or complementary contributions of local effort are easy to elicit, communities may be more likely to vote for politicians and parties that are most able to provide these goods. Uncooperative relations between groups, on the other hand, may result in each group voting for candidates who target government spending towards them. There is evidence, both for India and other parts of the world, of politicians transferring resources disproportionately towards the social groups to which they belong.[33] In this section we use data from several state level elections in the post independence period to examine whether the patterns of political participation vary systematically with social composition in a manner which is consistent with this hypothesis.

We focus primarily on 3 political variables- the total number of contestants, the share of votes received by the winning party, and a measure of political fragmentation which reflects the relative share of votes received by contesting parties. Our measure of political fragmentation is constructed in exactly the same way as our social fragmentation measure,using shares of total votes by each party instead of population shares of each social group. There are over 3000 state assembly constituencies in India, implying multiple constituencies in each district. The boundaries of assembly constituencies were drawn to make the number of voters in each constituency roughly equal and resulting in more constituencies in densely populated districts. To construct district level variables, we take averages across the district's constituencies. We also generate voter turnouts for each district (by summing voters and eligible voters for the district's constituencies).We have not have a prior prediction on how we expect turnouts to vary, since they are likely to depend both on individual propensities to engage in political activity and on the closeness of electoral contests.State assembly elections are usually held every five years, though not always in the same year for all states. To minimize the impact of year specific unobservable factors, we construct and use averages of our political variables for three time periods- the seventies, eighties and nineties.[34].

Table 4A, 4B, 4C and 4D presents results from least squares regressions of these political variables on social fragmentation and other control variables. Districts with more social fragmentation tend to have larger numbers of contestants

[33] Pande ()

[34] 1994 is the last election year in out dataset.

in their constituencies, more political fragmentation and smaller vote shares for the winning party. This is consistent with group based voting. Although the coefficient on social fragmentation is fairly consistent in sign across years and alternative measures of political fragmentation, it is not always statistically significant at conventional levels. It is therefore difficult to be confident that effect of heterogeneity operates entirely through its effect on political behavior.

7 To Conclude: The effect on the quality of life

In conclusion we briefly look at teh effect on some measures of quality of life. The above results would be uninteresting if there was absolutely no parallel effect on measures of the quality of life. A detailed investigation of thsi question is a subject of adifferent paper but here we briefly look at three indicators of the quality of life: literacy rates, infant mortality rates and crime rates. In a linear regression of literacy rates in 1991 on fragmentation and all our controls, the coefficient on fragmentation is -.18 for male literacy, -.06 for female literacy and -.11 for total literacy. These coeffcient are reported in Table 5A.

Running the same with infant mortality rates (defined per thousand live births), the coefficient is 55. If we use numbers that die before age 2 and 5 instead, the same picture emerges, the coefficient is 63 for under 2 and 122.9 for under five.

Finally if higher fragmentation is associated with more group conflict, we might also expect higher crime rates in more fragmented districts. We estimate least squares regressions, using the incidence of different types of crime at the district level as our dependent variable and our fragmentation measure, land inequality and the share of minority and religious groups as controls. Fragmentation increases crime by almost every measure.

References

[1] Alesina, Alberto, Reza Baqir and William Easterly (1999): "Public Goods and Ethnic Divisions", *Quarterly-Journal-of-Economics*; 114(4), pages 1243-84.

[2] Bates, Robert H. (1973): *Ethnicity in Contemporary Africa*, Monograph published by the Program of Eastern African Studies, Syracuse University.

[3] Beteille, Andre (1996): *Caste, Class and Power: Changing Patterns of Stratification in a Tanjore Village*, University of California Press.

[4] Chaudhur, Pramit (1979): *The Indian Economy: Poverty and Development*, Crosby Lockwood Staples.

[5] Dasgupta, Partha (1993): *An Inquiry into Well-Being and Destitution,* Oxford University Press.

[6] Duflo, Esther (2000): *Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from and unusual policy experiment*, working paper, department of economics, MIT.

[7] Easterly, William and Ross Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions", *Quarterly-Journal-of-Economics*; 112(4), pages 1203-50.

[8] Epstein, T. Scarlett, A.P. Suryanarayana and T. Thimegowda (1998): *Village Voices: Forty Years of Rural Transformation in South India*, Sage Publications, New Delhi.

[9] Esteban Joan and Debraj Ray (1999): "Conflict and Distribution" *Journal of Economic Theory*, 87, pages 379-415.

[10] Khwaja, Asim Ijaz (2000): *Can good projects succeed in bad communities? Collective Action in the Himalayas*, working paper, department of economics, Harvard University.

[11] Putnam, Robert D. (1993): *Making democracy work: civic traditions in modern Italy*, Princeton University Press.

[12] Shultz, T. Paul, (2001): *What are the returns to Poverty Reduction: Evaluation of Progresa in Mexico*, working paper, department of economics, Yale University.

[13] Singh, V.B. (1993): *Political Fragmentation and Electoral Process: 1991 Election in Uttar Pradesh*, Ajanta.

[14] Srinivas, M.N. (1996): *Indian Society through Personal Writings*, Oxford University Press.

[15] Wiener (1990): *The Child and the State in India*, Princeton University Press.

Table 0A: Descriptive Statistics
1981 public goods

Variable	Obs	Mean	Std. Dev.	Min	Max
anyed 81	356	0.7515886	0.1943949	0.1578947	1
primary 81	364	0.7367215	0.2016783	0.0431219	1
middle81	364	0.2128739	0.1958239	0	0.9913043
high81	364	0.0986634	0.1539319	0	0.9381443
juniorcollege81	364	0.0101824	0.0174001	0	0.1860465
anyed	356	0.7515886	0.1943949	0.1578947	1
primary	364	0.7367215	0.2016783	0.0431219	1
middle	364	0.2128739	0.1958239	0	0.9913043
high	364	0.0986634	0.1539319	0	0.9381443
juniorcoll~e	364	0.0101824	0.0174001	0	0.1860465
college	364	0.003971	0.015165	0	0.1486486
adltcen	364	0.060052	0.1220689	0	0.8571429
anymed	356	0.2158338	0.2068861	0.0183099	1
phs	364	0.0427393	0.0821089	0	1
phc	364	0.0333329	0.0659108	0	0.8571429
disp	364	0.0753993	0.1293525	0	0.8243243
hospital	364	0.0333143	0.1009632	0	1.195876
mcwcen	364	0.0398406	0.1029996	0	0.9302326
familyplanningcenter	364	0.037589	0.1056108	0	1.2
rp	363	0.0496715	0.0760214	0	0.5416667
smp	363	0.0038546	0.007807	0	0.0933333
childhealtwoker	364	0.0588573	0.1407865	0	0.8326119
anywater	356	0.9983929	0.0386109	0.8526316	1.672269
well	364	0.7271891	0.3473747	0	1.006896
twhp	364	0.1622371	0.2808574	0	1
tap	364	0.1407555	0.1977733	0	0.9988145
tank	364	0.137608	0.1936904	0	0.8571429
canal	364	0.0448081	0.1046334	0	0.8180649
fountain	364	0.0154434	0.0589429	0	0.7294118
anypower	364	0.5225077	0.3114459	0.0036364	1
powdomes	363	0.4310468	0.3427299	0.0036364	2.142857
powagric	363	0.3772099	0.3156197	0	1.545455
anypnt	356	0.2947599	0.2366183	0.0194805	1
post	364	0.21461	0.1436524	0	0.696868
phone	364	0.0256269	0.0562167	0	0.6666667
anycom	356	0.3498032	0.2743461	0.0072727	1
bus	364	0.3389679	0.2748358	0	1
rail	364	0.0178907	0.0251353	0	0.1532033
proad	364	0.3904718	0.2674059	0	1

Table 0B: Descriptive Statistics
1981 demographics and crime

Variable	Obs	Mean	Std. Dev.	Min	Max
tpop	377	1305908	1080912	18825	1.33E+07
psc	377	0.1554224	0.0972026	0	0.5259554
pst	377	0.154603	0.2545304	0	0.9809846
lratet	331	0.4051009	1.253337	0.0192	23.01
phindu	377	0.7854367	0.2691914	0.0130157	0.9981605
pmusli	377	0.096586	0.1647308	0.0003551	0.981764
pcris	377	0.0433705	0.1481876	0	0.9809659
psikh	377	0.0280643	0.1301745	0	0.9136912
brahman	366	0.0596299	0.0524876	0	0.3360027
crfrag	360	0.8468777	0.1918626	0.0361392	0.998234
ginia	366	0.691402	0.1723092	0.1700063	0.999446
murder	361	53.34488	43.89963	0	267
homicide	361	7.236842	12.77003	0	72
rape	361	12.2133	12.28957	0	131
kidnap	361	28.58864	31.32968	0	300
dacoity	361	35.47507	55.677	0	299
robbery	361	50.66759	58.71898	0	358
burglary	361	373.295	313.6551	2	2750
theft	361	860.3089	1061.455	1	12114.5
riots	361	250.8172	305.7015	0	3119
breach	361	41.68975	40.98646	0	320
cheating	361	32.29501	31.47681	0	269
counterfeit	361	1.790859	14.10122	0	253
totalcrime	361	2905.355	2341.407	12	20428

Table 0C: Descriptive Statistics
1991 public goods

Variable	Obs	Mean	Std. Dev.	Min	Max
anyed	445	0.8051185	0.1656742	0.2284382	1
primary	445	0.7859392	0.175482	0.2132867	1
middle	445	0.258723	0.1995477	0	1
high	445	0.1285234	0.1663027	0	1
juniorcollege	445	0.019132	0.0315111	0	0.3529412
college	445	0.0050109	0.0124364	0	0.1411765
adltcen	445	0.0917461	0.1411335	0	0.9146342
anymed	445	0.3828726	0.2987728	0.0345679	1
phs	445	0.1014284	0.1281404	0	1
phc	445	0.0519772	0.0845995	0	0.7142857
hcen	445	0.0126796	0.0201488	0	0.1647059
disp	445	0.06252	0.0981441	0	0.6206896
hospital	445	0.0267503	0.0486198	0	0.4456522
nursinghome	445	0.0185166	0.0964805	0	0.7888889
mcwcen	445	0.0290678	0.0584365	0	0.5411765
maternityhome	445	0.0097498	0.0238884	0	0.2715232
childwelfarecenter	445	0.0221298	0.043652	0	0.4245283
family planning center	445	0.0306263	0.0688818	0	0.4507042
tbclinics	445	0.0017809	0.0077671	0	0.1045455
rp	445	0.0867914	0.1151702	0	0.72
smp	445	0.0051578	0.0163197	0	0.2616279
childhealthworker	445	0.1586911	0.2667398	0	0.9991357
anywater	445	0.9978223	0.0061789	0.9284512	1.00004
well	445	0.6717745	0.3345141	0	1
handpump	445	0.530789	0.3875528	0	1
tubewell	445	0.2137822	0.2884659	0	0.9963099
tap	445	0.2831007	0.3015061	0	1
tank	445	0.1451604	0.2135676	0	0.9721362
canal	445	0.0386272	0.0828667	0	0.8760703
nala	445	0.0484457	0.1206937	0	0.8
fountain	445	0.022333	0.0759038	0	0.8955512
anypower	445	0.7574677	0.2623224	0.0654634	1
powdomes	445	0.7105089	0.2799253	0.0436364	1
powagric	445	0.5118226	0.3691621	0	1
allpower	445	0.3211254	0.3107049	0	1

Table 0D: Descriptive Statistics
1991 public goods continued

Variable	Obs	Mean	Std. Dev.	Min	Max
anypnt	445	0.3212323	0.2506258	0.010274	1
post	445	0.291538	0.2341491	0.010274	1
tgraph	445	0.0491346	0.1154289	0	0.8571429
phone	445	0.0985402	0.1598194	0	1
anycom	445	0.4379819	0.2839129	0.0345238	1
bus	445	0.4277001	0.284979	0	1
rail	445	0.0175978	0.0236173	0	0.2
proad	445	0.4511706	0.2724771	0	1

Table 0E: Descriptive Statistics
1991 demographics and crime

Variable	Obs	Mean	Std. Dev.	Min	Max
psc	445	0.1598418	0.094514	0	0.5450974
pst	445	0.1686104	0.2651369	0	0.9849431
lratem	445	0.4799459	0.1246507	0.1558354	0.8633283
lratef	445	0.2613191	0.1531961	0.0333869	0.8388228
lratet	445	0.3746591	0.1343338	0.1042462	0.8510541
phindu	445	0.7888431	0.2524504	0.0122839	0.9975895
pmusli	445	0.0890628	0.1251593	0.0003791	0.9606515
pcris	445	0.0623255	0.1879927	4.70E-06	0.9863827
psikh	437	0.0250096	0.1250347	1.02E-06	0.9395128
brahman	436	0.0577394	0.0517578	0	0.3356538
crfrag	429	0.8464384	0.1845375	0.026855	0.9983052
crfragsq	429	0.7504326	0.2266333	0.0007212	0.9966133
ginia	403	0.6378725	0.1264129	0.1276563	0.8503081
murder	438	73.76712	66.08625	0	450.5
attemptedmurder	438	55.78082	74.16458	0	692
homicide	438	9.069635	15.79956	0	114.5
rape	438	21.24087	19.38434	0	135.5
kidnap	438	37.14155	40.03212	0	386
womenkidnap	438	27.71918	29.57278	0	183.5
otherskidnap	438	9.442922	21.81433	0	305
dacoity	438	22.80023	31.49161	0	339
dacoityprep	438	2.744292	6.966553	0	83.5
robbery	438	49.90753	58.497	0	441
burglary	438	259.6781	231.0747	0	1780
theft	438	607.6792	691.6693	3.5	7117
riots	438	218.1678	247.2514	0	2303
breach	438	30.06621	35.82434	0	273
cheating	438	43.99315	59.84864	0	725
counterfeit	438	4.939498	37.31841	0	766
othercrime	438	1725.525	1513.644	8	11701
totalcrime	438	3162.326	2595.046	22	23125

Table 1A: Education and Health Facilities

public good	Estimated Coefficients and Standard Errors							
	Basic specification		Controlling for average population density and bank deposits per capita					
	(1)	(2)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
	frag[a] 1981	frag[a] 1991	frag[b] 1981	ST 1981	gini 1981	frag[b] 1991	ST 1991	gini 1991
any educational facility	-0.62	-0.12	-0.72	0.32	-0.15			
	(0.59)	(0.58)	(0.56)	(0.32)	(0.63)			
primary schools	-0.44	-0.33	-0.53	0.22	-0.05	-0.78	0.39	0.16
	(0.60)	(0.57)	(0.58)	(0.32)	(0.64)	(0.56)	(0.34)	(0.46)
middle schools	-1.43 **	-1.53 **	-1.38 **	-0.79 **	-0.14	-1.79 **	-0.87 **	0.13
	(0.47)	(0.44)	(0.44)	(0.27)	(0.47)	(0.42)	(0.25)	(0.37)
high schools	-2.21 **	-1.56 **	-2.21 **	-1.01 **	0.66	-1.68 **	-0.80 **	0.39
	(0.51)	(0.44)	(0.48)	(0.30)	(0.56)	(0.42)	(0.25)	(0.45)
junior colleges	-0.05	-0.48	-0.66	-0.53 *	1.08	-0.93 *	-0.79 **	0.78 *
	(0.65)	(0.52)	(0.61)	(0.31)	(0.72)	(0.51)	(0.29)	(0.43)
colleges	-1.98 *	-0.05	-1.56	-0.67	0.07	0.98	-1.22 **	-0.28
	(1.08)	(0.94)	(0.99)	(0.60)	(0.94)	(0.90)	(0.49)	(0.65)
adult literacy centres	-0.99	0.12				-0.26	0.62	1.28 *
	(0.82)	(0.94)				(0.99)	(0.45)	(0.77)
any medical facility	-1.01 *	0.90	-0.84 *	-0.61 **	-0.45	0.81	-0.38	-0.05
	(0.51)	(0.57)	(0.49)	(0.29)	(0.55)	(0.53)	(0.33)	(0.53)
primary heatlh subcentre	-1.67 **	0.41	-1.75 **	-0.91 **	0.72	0.54	-0.16	-0.25
	(0.58)	(0.64)	(0.59)	(0.27)	(0.54)	(0.66)	(0.32)	(0.55)
primary health centre	-1.00 **	-0.47	-0.72	-0.33	-0.04	-0.24	-0.35	-0.67 *
	(0.50)	(0.44)	(0.50)	(0.26)	(0.47)	(0.43)	(0.22)	(0.37)
health centre		-1.50 **				-0.79	0.21	0.69
		(0.75)				(0.73)	(0.37)	(0.58)
dispensary	-1.75 **	-1.30 **	-1.61 **	-0.58 **	-0.09	-0.83 *	-0.51 **	-0.15
	(0.45)	(0.49)	(0.47)	(0.26)	(0.52)	(0.48)	(0.26)	(0.47)
hospital	-2.72 **	-2.56 **	-2.74 **	-0.02	0.86	-2.53 **	-0.24	0.84 *
	(0.58)	(0.52)	(0.56)	(0.34)	(0.60)	(0.50)	(0.30)	(0.47)
nursing home		-1.54 *				0.03	-0.56	-0.15
		(0.84)				(0.80)	(0.51)	(0.76)
maternity and child welfare centre	0.18	2.03 **	0.16	-0.08	0.89	2.06 **	1.51 **	0.68
	(0.65)	(0.77)	(0.64)	(0.33)	(0.68)	(0.79)	(0.51)	(0.63)
maternity home		-2.30				-2.56 *	-0.28	6.72 **
		(1.44)				(1.37)	(0.56)	(1.07)
child welfare centre		3.11 **				2.81 **	-0.07	2.26 **
		(0.94)				(0.90)	(0.49)	(0.81)
family planning centre	-0.64	1.60	-0.60	-0.54 *	1.16	1.62 *	-0.83 **	2.50 **
	(0.77)	(1.00)	(0.80)	(0.30)	(0.71)	(0.98)	(0.34)	(0.67)
tubercolosis clinic		-1.91				-1.22	2.34 **	-3.23 **
		(1.32)				(1.43)	(0.60)	(1.55)
subsidized medical practitionner	-0.47	2.70 **	-0.80	-0.83	-0.98	2.06 **	-0.25	0.20
	(1.11)	(0.72)	(1.09)	(0.54)	(0.94)	(0.82)	(0.52)	(0.82)
child health worker	-3.75 **	2.77 **	-3.61 **	-0.40	-0.87	3.10 **	-0.01	-1.61 *
	(1.10)	(0.94)	(1.24)	(0.54)	(1.06)	(0.97)	(0.61)	(0.87)
registered medical practitioners	**-0.98**	**0.28**	**-0.28**	**-1.77 ****	**-0.23**	**0.20**	**-2.44 ****	**1.99 ****
	(0.78)	**(0.89)**	**(0.76)**	**(0.52)**	**(0.87)**	**(0.79)**	**(0.45)**	**(0.67)**

Table 1B: Water, Power and Communication Facilities

public good	Estimated Coefficients and Standard Errors							
	Basic specification		Controlling for average population density and bank deposits per capita					
	(1)	(2)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
	frag[a] 1981	frag[a] 1991	frag[b] 1981	ST 1981	gini 1981	frag[b] 1991	ST 1991	gini 1991
any water	-2.39	0.71	-6.52 **	1.47	5.57 **	-0.17	1.01	1.52
	(2.79)	(1.52)	(2.72)	(1.04)	(1.72)	(1.72)	(1.05)	(1.25)
wells	4.80 **	2.57 **	4.72 **	-1.53 **	3.30 **	3.10 **	-0.08	0.09
	(1.26)	(1.25)	(1.36)	(0.62)	(1.31)	(1.27)	(0.57)	(0.96)
handpumps		0.55				1.09	0.48	1.22
		(1.29)				(1.31)	(0.52)	(0.87)
tubewells		-0.14				0.71	-2.34 **	-1.60 **
		(0.97)				(1.00)	(0.51)	(0.73)
tubewells/handpumps	-4.63 **		-2.93 **	-3.36 **	-0.98			
	(1.24)		(1.33)	(0.70)	(1.25)			
tapped water	-2.63 **	-4.21 **	-2.65 **	-0.63	-0.59	-4.41 **	-1.11 **	0.20
	(0.67)	(0.79)	(0.71)	(0.50)	(1.10)	(0.85)	(0.51)	(0.81)
tanks	-0.56	-0.40	1.13	-0.93 *	-2.42 **	0.64	-0.76	-2.19 **
	(1.34)	(1.53)	(1.47)	(0.53)	(0.92)	(1.61)	(0.54)	(0.93)
canals	-2.10	4.36 **	-2.66 *	0.05	1.58	5.20 **	0.70	0.18
	(1.32)	(1.08)	(1.38)	(0.76)	(1.22)	(1.13)	(0.53)	(0.94)
fountains	-0.94	-3.79 **	0.87	1.97 *	-2.74	-3.44 **	1.16	-1.57
	(2.31)	(1.46)	(2.11)	(1.10)	(2.77)	(1.42)	(1.20)	(1.55)
any power	1.04	-3.81 **	0.75	-1.35 **	1.26	-1.42 *	-1.05 **	0.96 *
	(0.85)	(1.48)	(0.83)	(0.43)	(0.85)	(0.75)	(0.41)	(0.55)
power for domestic use	-0.60	-1.58 **	-1.19 *	-1.39 **	2.05 **	-1.71 **	-0.89 **	0.73
	(0.74)	(0.68)	(0.72)	(0.40)	(0.78)	(0.69)	(0.38)	(0.51)
power for agricultural use	0.08	0.76	-0.33	-1.46 **	1.16	0.64	-0.84 *	0.77
	(1.12)	(1.15)	(1.16)	(0.49)	(0.96)	(1.15)	(0.50)	(0.71)
power for all uses		0.16				-0.38	-1.55 **	1.62 **
		(0.84)				(0.86)	(0.44)	(0.67)
any post or telegraph facility		-1.29 **				-1.54 **	-0.72 **	0.30
		(0.58)				(0.58)	(0.34)	(0.52)
post offices	-1.22	-1.28 **	-1.21 **	-0.41 *	-0.17	-1.59 **	-0.48	0.30
	(0.43)	(0.59)	(0.42)	(0.25)	(0.47)	(0.60)	(0.34)	(0.53)
telgraph offices		-0.84				-1.13 *	-1.12 **	1.36 **
		(0.63)				(0.62)	(0.34)	(0.57)
phone connections	-0.12	-0.40	-0.14	-1.34 **	2.49 **	-0.72	-1.67 **	1.58 **
	(0.86)	(0.64)	(0.84)	(0.46)	(0.80)	(0.61)	(0.31)	(0.59)
any communication facility	-1.07	-0.61	-0.96 *	-0.19	-0.61	-0.91 *	-0.66 **	0.33
	(0.51)	(0.50)	(0.50)	(0.29)	(0.55)	(0.51)	(0.30)	(0.45)
bus services	-1.08	-0.64	-0.99 *	-0.25	-0.56	-0.95 *	-0.64 **	0.36
	(0.52)	(0.50)	(0.51)	(0.30)	(0.56)	(0.51)	(0.30)	(0.45)
rail services	-3.48	-0.28	-3.68 **	-1.10 **	2.30 **	0.13	-0.89 **	-0.65
	(0.73)	(0.70)	(0.76)	(0.49)	(0.92)	(0.71)	(0.35)	(0.56)
paved roads	1.19	1.32 **	1.17 **	0.03	-0.05	0.78 *	-0.39	0.81 **
	(0.49)	(0.49)	(0.46)	(0.24)	(0.49)	(0.47)	(0.24)	(0.37)

Table 2A: Education and Health Facilities

public good	Estimated Coefficients and Standard Errors			
	Linear Probability Model		Controlling for registered medical practitioner	
	frag 1981	frag[a] 1991	frag[b] 1981	frag 1991
	(1)	(2)	(3)	(4)
any educational facility	0.00	-0.03	-0.65	-0.60
	(0.10)	(0.09)	(0.53)	(0.56)
primary schools	0.03	-0.03	-0.44	-0.78
	(0.11)	(0.09)	(0.55)	(0.55)
middle schools	-0.16 **	-0.27 **	-1.06 **	-1.58 **
	(0.07)	(0.08)	(0.40)	(0.39)
high schools	-0.15 **	-0.18 **	-2.10 **	-1.66 **
	(0.05)	(0.05)	(0.46)	(0.40)
junior colleges	-0.01	-0.02 **	-0.71	-0.95 *
	(0.01)	(0.01)	(0.60)	(0.50)
colleges	-0.01	-0.01	-1.27	0.98
	(0.01)	(0.01)	(0.99)	(0.86)
adult literacy centres	0.00	-0.02	-0.17	-0.10
	(0.06)	(0.06)	(0.84)	(0.99)
any medical facility	-0.09	-0.40 **	-0.53	0.83
	(0.09)	(0.17)	(0.46)	(0.51)
primary heatlh subcentre	-0.08 **	-0.01	-1.59 **	0.61
	(0.04)	(0.06)	(0.59)	(0.66)
primary health centre	-0.02 **	-0.02	-0.68	-0.21
	(0.01)	(0.02)	(0.50)	(0.42)
health centre		-0.02 **		-0.88
		(0.01)		(0.73)
dispensary	-0.06	-0.05	-1.42 **	-0.88 *
	(0.04)	(0.04)	(0.46)	(0.48)
hospital	-0.01	-0.07 **	-2.69 **	-2.54 **
	(0.04)	(0.02)	(0.55)	(0.50)
nursing home		-0.04 **		0.05
		(0.02)		(0.80)
maternity and child welfare centre	0.06	0.03	0.71	2.34 **
	(0.04)	(0.02)	(0.59)	(0.79)
maternity home		0.00		-2.47 *
		(0.01)		(1.37)
child welfare centre		0.03		2.60 **
		(0.03)		(0.91)
family planning centre	-0.05 **	0.01	-0.54	1.44
	(0.02)	(0.02)	(0.79)	(0.95)
tubercolosis clinic		0.00		-1.55
		(0.01)		(1.32)
subsidized medical practitionner	0.00	0.02	-0.34	1.92 **
	(0.00)	(0.02)	(1.07)	(0.77)
child health worker	-0.12	0.51 **	-3.47 **	3.04 **
	(0.08)	(0.16)	(1.27)	(0.97)

Table 2B: Water, Power and Communication Facilities

public good	Estimated Coefficients and Standard Errors			
	Linear Probability Model		Controlling for registered medical practitioners	
	frag[a] 1981	frag[a] 1991	frag[b] 1981	frag 1991
	(1)	(2)	(3)	(4)
any water	0.00	0.01	-4.83 *	0.06
	(0.01)	(0.01)	(2.94)	(1.72)
wells	0.74 **	0.77 **	4.74 **	2.99 **
	(0.26)	(0.26)	(1.37)	(1.28)
handpumps		0.88 **		1.13
		(0.24)		(1.31)
tubewells		0.07		0.85
		(0.09)		(0.99)
tubewells/handpumps	-0.05		-2.87 **	
	(0.09)		(1.34)	
tapped water	-0.32 **	-0.75 **	-2.66 **	-4.42 **
	(0.13)	(0.20)	(0.71)	(0.85)
tanks	0.09	0.07	0.76	0.66
	(0.07)	(0.07)	(1.48)	(1.60)
canals	0.01	0.09	-2.78 **	5.18 **
	(0.12)	(0.11)	(1.39)	(1.13)
fountains	-0.06	-0.24 **	1.32	-3.21 **
	(0.05)	(0.07)	(2.11)	(1.35)
any power	0.16	-0.13	1.03	-1.46 **
	(0.12)	(0.10)	(0.80)	(0.73)
power for domestic use	-0.02	-0.23 **	-0.91	-1.69 **
	(0.14)	(0.10)	(0.69)	(0.67)
power for agricultural use	0.33 *	0.36 *	0.05	0.73
	(0.17)	(0.19)	(1.14)	(1.14)
power for all uses		0.06		0.02
		(0.12)		(0.84)
any post or telegraph facility		-0.25 *		-1.37 **
		(0.13)		(0.56)
post offices	-0.09	-0.28 **	-1.04 **	-1.41 **
	(0.08)	(0.13)	(0.40)	(0.58)
telgraph offices		-0.04		-1.02 *
		(0.03)		(0.60)
phone connections	-0.02	0.01	0.27	-0.57
	(0.02)	(0.07)	(0.82)	(0.58)
any communication facility	-0.16 *	-0.21 **	-0.69	-0.83 *
	(0.09)	(0.09)	(0.47)	(0.48)
bus services	-0.09	-0.21 **	-0.74	-0.87 *
	(0.10)	(0.09)	(0.48)	(0.48)
rail services	0.01	0.00	-3.32 **	0.17
	(0.02)	(0.01)	(0.75)	(0.70)
paved roads	0.18 *	0.13	1.35 **	0.97 **
	(0.10)	(0.10)	(0.45)	(0.45)

Table 3A: Magnitude of the Effect

Variable	Mean	year	coefficient	marginal effects
middle	0.2128739	1981	-1.38	-0.231230872
high	0.0986634	1981	-2.21	-0.196532943
anymed	0.2158338	1981	-0.84	-0.142169639
phs	0.0427393	1981	-1.74	-0.071188015
disp	0.0753993	1981	-1.61	-0.112239935
hospital	0.0333143	1981	-2.74	-0.088240213
childhealthworker	0.0588573	1981	-3.6	-0.199415226
anywater	0.9983929	1981	-6.5	-0.010429362
well	0.7271891	1981	4.7	0.93241003
tap	0.1407555	1981	-2.65	-0.320499981
canal	0.0448081	1981	-2.65	-0.113420886
powdomestic	0.4310468	1981	-1.19	-0.291842093
post	0.21461	1981	-1.2	-0.202263057
anycom	0.3498032	1981	-0.96	-0.218343284
bus	0.3389679	1981	-0.99	-0.221827976
rail	0.0178907	1981	-3.68	-0.064659892
proad	0.3904718	1981	1.17	0.278464181
twhp	0.1622371	1981	-2.92	-0.396875372

Table 3B: Magnitude of the Effect

Variable	Mean	year	coefficient	marginal effects
middle	0.258723	1991	-1.79	-0.343295883
high	0.1285234	1991	-1.68	-0.188168628
juniorcoll~e	0.019132	1991	-0.93	-0.017452349
disp	0.06252	1991	-0.84	-0.04923345
hospital	0.0267503	1991	-2.53	-0.065867845
mcwcen	0.0290678	1991	2.06	0.058139098
maternityhome	0.0097498	1991	-2.5	-0.024136853
childwelfarecenter	0.0221298	1991	2.8	0.060592201
fpc	0.0306263	1991	1.62	0.048095094
smp	0.0051578	1991	2.51	0.012879305
childhealtworker	0.1586911	1991	3.1	0.413875528
well	0.6717745	1991	3.1	0.683529916
tap	0.2831007	1991	-4.4	-0.893000652
fountain	0.022333	1991	-3.44	-0.075109776
anypower	0.7574677	1991	-1.41	-0.259031641
powdomes	0.7105089	1991	-1.71	-0.351723065
anypnt	0.3212323	1991	-1.54	-0.335784849
post	0.291538	1991	-1.59	-0.328404315
tgraph	0.0491346	1991	-1.12	-0.052326838
anycom	0.4379819	1991	-0.91	-0.223999917
bus	0.4277001	1991	-0.95	-0.232534088
proad	0.4511706	1991	0.78	0.193140238

Table 4A: Political Behavior Dependent Variable is Number of Contesting Candidates

	1971	1981	1991
fragment	.95*	.30	2.0
	(.57)	(.41)	(1.37)
gini coefficient	.19	.19	.36
	(.49)	(.42)	(1.02)
scheduled castes	.11	.62	4.0***
	(.51)	(.52)	(1.48)
scheduled tribes	-.84**	-.96**	-.79
	(.35)	(.33)	(.64)
state dummies	yes	yes	yes
observations	252	306	293
R^2	.50	.65	.73

Table 4B: Political Behavior Dependent Variable is Political Fragmentation

	1971	1981	1991
fragment	.07	.10**	.05
	(.05)	(.04)	(.04)
gini coefficient	-.04	-.07*	.01
	(.04)	(.04)	(.04)
scheduled castes	-.02	.05	.08
	(.03)	(.05)	(.05)
scheduled tribes	-.035*	-.05*	.001
	(.02)	(.028)	(.024)
state dummies	yes	yes	yes
observations	252	306	293
R^2	.51	.64	.63

Table 4C: Political Behavior Dependent Variable is Share of the party with maximum votes

	1971	1981	1991
fragment	-.11*	-.13***	-.05
	(.06)	(.04)	(.04)
gini coefficient	.06	.09**	-.03
	(.04)	(.04)	(.04)
scheduled castes	.03	-.05	-.06
	(.04)	(.05)	(.05)
scheduled tribes	.054***	.09***	.01
	(.019)	(.03)	(.03)
state dummies	yes	yes	yes
observations	252	305	293
R^2	.48	.69	.59

Table 4D: Political Behavior: Dependent Variable is Voter Turnouts

	1971	1981	1991
fragment	.11 (.09)	.097 (.053)	.06 (.04)
gini coefficient	.08 (.05)	.01 (.05)	.05 (.04)
scheduled castes	.05 (.06)	.04 (.07)	-.02 (.07)
scheduled tribes	-.15*** (.04)	-.19*** (.03)	-.18*** (.03)
state dummies	yes	yes	yes
observations R^2	252 .65	306 .78	293 .83

Notes: other controls included are the same as in tables 1 and 2, column (b). Values for per capita bank deposits were not available for 1971, and 1981 values were used instead.

Table 5A Effect on literacy in 1991
Estimated Coefficients and Standard Errors

	frag	ST	gini
Male literacy	-0.19 **	-0.18 **	0.11 **
	(0.06)	(0.04)	(0.05)
Female literacy	-0.07	-0.12 **	0.05
	(0.06)	(0.04)	(0.05)
Total literacy	-0.12 **	-0.15 **	0.09 *
	(0.06)	(0.03)	(0.05)

Table 5B: Fragmentation and Crime

Type of Crime	1981	1991
murder	70.7**	52.5
	(32.1)	(35.7)
rape	15.9***	16.3**
	(5.6)	(7.9)
kidnap	36.0***	43.6
	(14.2)	(21.3)**
burglarly	400.3**	88.4
	(159.8)	(100.3)
crimes against the person	158.5***	198.4**
	(51.3)	(85.6)
crimes against property-violent	247***	111.6***
	(74.2)	(40.3)
crimes against property-non violent	1293.9	472.4
	(793.1)	(487)
mean R^2	.48	.41
mean observations	317	363