History, Social Divisions and Public Goods in Rural India *

Abhijit Banerjee†
Lakshmi Iyer‡
Rohini Somanathan§

Abstract

We examine the influence of three historically important sources of social divisions on the availability of public goods in rural India: colonial power, landowner-peasant relations as determined by the land tenure system and social fragmentation based on the Hindu caste system and the presence of sizable religious minorities. Using data on public goods from 1991, we find that regions that were under British colonial power in the pre-independence period and those where agrarian power was concentrated in the hands of landlords have lower access to these goods as do areas with high levels of social fragmentation. (JEL: H41, P16)

*We thank participants at the European Economic Association Meeting for helpful comments and suggestions.
†Department of Economics, MIT. banerjee@mit.edu
‡Harvard Business School. liyer@hbs.edu
§Department of Economics, University of Michigan. rohinis@umich.edu

1 Introduction

One of the most powerful hypotheses in political economy is the notion that social divisions undermine economic progress, not just *in extremis*, as in the case of a civil war, but also in more normal times. The idea is that even in a market economy there are numerous transactions that rely on the cooperation of others: trade often requires trust, providing public goods needs collective action and the rule of law is only possible if everyone accepts the rights of others. Homogeneous societies, it is suggested, have an advantage because there may be more contact across the population, which builds understanding, trust and empathy, and shared interests. This makes it more likely that they will all be on the same side. If true, this hypothesis has a number of important implications. Among these is an argument for a high degree of sensitivity to distributional issues, since the memory, real or imagined, of having been exploited, can create a divide that will continue to hurt the economy many years into the future.

Unfortunately it is difficult to measure social divisions directly. One strategy is to look for social and political facts of history that make certain places more prone to social conflict than others. This is the approach taken by Easterly and Levine (1997) when they proxy social divisions in African countries by the ethno-linguistic fragmentation of the population, and by Alesina, Baqir and Easterly (1999) when they look at ethno-linguistic fragmentation and the provision of public goods in US cities, as well as by a substantial recent literature in the same tradition.[1]

We focus on a large set of public goods in rural India in 1991 and use three variables capturing historically important institutional sources of differences in social inequalities. Our approach is similar to that taken in the existing literature, but the nature of our data allows us to take this agenda considerably further. In particular, we are able to compare the effects of caste and class divisions, and the historical variation also allows us to construct credible instruments for some of these sources of social inequities.

Our first variable is a dummy for whether the district was under direct British rule or whether it was a part of a "native state", formally a protectorate of the British with autonomy on internal matters. The British conquest of India began in 1757 and ended in 1858, when, faced with a major rebellion in India, the British Crown took over the government of India from the East India Company. Queen Victoria then assured the rulers of the hitherto unconquered parts of India that she would respect their rights over their kingdom, as long as they accepted British sovereignty. These "native states" retained considerable internal autonomy till the British left in 1947. At that time, around 560 big and small native states were absorbed into independent India. The impact of

[1] See Miguel and Gugerty (2004), Khwaja (2002) and Bardhan and Dayton-Johnson (2002), among others.

this on social divisions could go either way: on the one hand, British rule in India was explicitly built on the principle of "divide and rule"—solidarity among Indians was discouraged. Moreover British rule led to a concentration of power in the metropolitan centers and an emphasis on an English education, which contributed to the creation of a new metropolitan-focussed Indian elite, cut off from the rural areas where most Indians lived. On the other hand, the British probably helped to modernize Indian society to some extent, for example by introducing a degree of meritocracy through the system of public exams. By contrast the rulers of the native states were clearly more invested in the traditional structures of authority, including the caste system. On balance therefore, it is not clear whether we should expect the British areas to do better or worse.

Our second variable captures variations in class structure based on the land revenue policies of the British colonial state: in some areas of British India, the responsibility for collecting land revenue was given to a landlord (the so-called *Zamindari* system), thereby setting up the basis for an ongoing conflict between peasants and landlords. Elsewhere, the colonial state directly collected the land revenue from the cultivator (the *raiyatwari* system), thereby avoiding this particular source of internecine conflict.

Our third variable is an index of caste and religion based fragmentation. This is in some ways a more reliable measure than the usual ethno-linguistic fragmentation index: caste and religious classifications in India are ancient and there are very explicit rules about who belongs to which group, which makes it less likely that ethnic identity is endogenous to the politico-economic environment. For the purpose of avoiding endogeneity, it also helps that our caste data is from the 1920s, long before caste became politicized and before the state took an active role in providing most public goods. Our premise is that caste and religious divisions create an environment where social relations can easily become conflictual; both caste and religious conflict have certainly been an important part of the Indian political landscape in the recent past.

The rest of the paper proceeds as follows: section 2 describes the institutional environment and the data, section 3 reports our results and section 4 discusses possible interpretations.

2 Institutional environment and data sources

We use district-level data from 391 districts in 16 major states of India. We focus on 26 distinct categories of public goods. The 1991 Census of India lists the district-wise fraction of villages with access to each of these goods. We exclude all facilities which are found, on average, in less that 2 percent of villages in a district on the grounds that these amenities are not typically located

in rural areas. Our final list of 26 goods is extremely diverse: it includes educational institutions, health care centers, a variety of water sources, availability of electric power, transportation facilities and communication infrastructure. The census data does not distinguish between state-funded and privately-funded facilities but during the period of our study, most of these were financed by the state.

Our data exhibits enormous variation in access to public goods within India: in 1991 rural Kerala had some medical facility in 96 percent of villages, as compared to 11 percent in rural Orissa. Within the state of Andhra Pradesh, less than 7 percent of the villages in Vishakhapatnam district had middle schools while 55 percent in Guntur had one. What makes these differences remarkable is that they have remained so large, despite the fact that the Indian government, for many years and at all levels, has been explicitly committed to equalizing access to public goods.[2] Prima facie, this suggests a role for the forces of political economy. We capture different aspects of this political economy in three variables described briefly above.

Most districts in 1991 were historically either fully under British rule or part of a princely state. For the few districts which are mixed, we chose the description that covered the majority of the land, to end up with a classification of all districts into a dummy that is zero for directly ruled and one for indirectly ruled districts. We refer to this as the "non-British" dummy; around one-third of the districts in our sample are classified as non-British. For the land tenure classification, we used historical records to determine the fraction of land that was not under the landlord system and this is the value taken by the "non-landlord" variable. The mean of this variable in the British districts is 0.52. We set the non-landlord variable to be zero in non-British areas. Therefore, in a regression which includes both the non-British and the non-landlord variables, the non-British areas are being compared to directly ruled landlord areas (the omitted category). The classification of Indian districts by colonial influence is based on Iyer (2003) and that in terms of the land tenure system is based on Banerjee and Iyer (2003).

The third element in the political economy is a measure of caste and religious differences, which have been important in India for a very long time. The caste data is based on the 1931 census (which is the last census which comprehensively recorded this data); we extrapolate this to the present-day population shares of Hindus in each district. We then construct a fragmentation index, along the lines of the standard ethno-linguistic fragmentation index, assuming that Muslims, Christians, Sikhs, Buddhists and Jains are homogenous groups but that Hindus are divided into 185 major caste groups (Banerjee and Somanathan (2001) describe this construction in greater detail). Indian

[2]See Banerjee and Somanathan (2004) for some examples of the kinds of commitments that were made in this regard.

districts are extremely diverse: on a scale of zero to one, the average fragmentation index in our sample is 0.88.

The regression we run is the following:

$$y_{ij} = \alpha_j + \beta_j NB_i + \gamma_j NL_i + \delta_j FRAG_i + X_i\Theta_j + \varepsilon_{ij} \quad (1)$$

where y_{ij} is the share of villages with public good j in district i, NB_i is the non-British dummy, NL_i is the fraction of land not under landlord control in district i, $FRAG_i$ is the fragmentation index, and X_i is a vector of other controls. Among the controls we include geographical factors which may affect economic conditions or the ease of delivering public goods: average annual rainfall, maximum and minimum temperatures, a dummy for coastal districts, the proportion of wasteland area which is barren or rocky, sandy or mountainous. We also include controls for average village population and the number of villages in the district, because it is probably easier to deliver public goods to concentrated populations. To avoid the fragmentation variable picking up differences in the demographic composition of the population, we control for the proportions of Scheduled Castes, Scheduled Tribes, Muslims, Christians, Sikhs and Brahmans in the population. Finally, we control for the Gini coefficient of land ownership to ensure that our non-landlord variable picks up the effect of the social divisions created by the empowerment of landlords, rather than the effect of differences in the pattern of land ownership today.

We use instrumental variables to control for the possible endogeneity of our non-British dummy and the non-landlord variable. Our instrument for the non-British variable with a dummy for whether the ruler of the native state died without a natural son in the period 1848-1856. During these years, the Governor-General Lord Dalhousie used the "Doctrine of Lapse" to annex these heirless states. As long as we accept that death without a natural son is a more or less random event, this is a valid instrument. Since this variable is only defined for areas that had not already been annexed before 1848, we include a dummy in all our regressions for whether the area was annexed before 1848. We deal with the potential endogeneity of the land revenue system by using the fact of being conquered by the British in the period 1820-1856 as an instrument for being non-landlord. Detailed discussions of the validity of these instruments can be found in Iyer (2003) and Banerjee and Iyer (2003), respectively.

3 Results

Table 1 summarizes the main results from estimating (1) using OLS and IV. Column 1 gives the mean fraction of villages in a district with a particular public good. Columns 2-4 report, respec-

tively, the OLS coefficients on the non-British, the non-landlord and the fragmentation variables. Columns 5-7 report the corresponding IV estimates. The first stage for the IV is not reported here, but the instruments come in strongly significant.

The non-British coefficient is significant and positive for 14 goods and negative and significant for 4 in the OLS. When we do the IV, it remains positive and significant in 9 and is now negative and significant in 3. The OLS and the IV coefficients are fairly similar in magnitude when they are both significant, and have the same sign. In most cases, the IV coefficient is slightly larger than the OLS.

The OLS estimates of the non-landlord effect are positive and significant for 15 of the 26 goods and negative and significant for 3. When we do IV, 7 are positive and significant and 2 negative and significant. Again, in the case where both the IV and the OLS are significant, they have the same sign, and the magnitudes are similar, with one exception (handpumps). When we compare the coefficients on the non-landlord and non-British variables we see that they are often similar in magnitude and when they are significant, almost always have the same sign (except in one case). This says that a district that was entirely non-landlord in British India did about as well as a non-British district, compared to a British landlord district.

Finally, the caste and religious fragmentation coefficient is negative and significant for 10 goods and positive and significant for 4 in the OLS. When we instrument for both the non-landlord and the non-British variables, it remains negative and significant for 10 goods and positive and significant for 3. The point estimates in all the cases where both are significant have the same sign and are remarkably similar in magnitude.

To compare the non-landlord coefficient with the fragmentation coefficient, we need to control for the higher variance of the former. When we compare the effects of a one standard deviation change in each, the results are not uniform: of the 8 cases where both variables are significant in the IV specification, 2 have a larger non-landlord effect, 2 have a larger fragmentation effect and 4 have very similar effects. Indian sociologists have often debated whether caste is more important than class in determining agrarian relations. Our results suggest the salience of both these divides.

It is also striking that some water facilities behave quite differently from public goods related to education, communications and electricity. For wells and handpumps, all three of our variables have the opposite sign. For these we observe a negative non-landlord and non-British effect, and a positive fragmentation effect and for tanks, the non-landlord and the fragmentation index both have the opposite sign. These facilities are often created by private parties and these results suggest that the proliferation of water sources is actually a symptom of social disunity; that people are willing to

share water sources in places with smaller social divisions, whereas in a more divisive environment, each group invests in getting its own supply. This is consistent with the enormous salience of the act of sharing water within the Indian sociocultural framework: in fact, anthropologists often use the willingness to accept a drink of water from a member of another caste as an index of the relation between the two castes.

4 Conclusions

The non-landlord and fragmentation results are consistent with our priors. The non-British result is also surprisingly clear-cut: British (landlord) areas do significantly worse than non-British areas. The historical nature of our data together with our instrumental variable approach makes us fairly confident of these observed relationships.[3]

The next step is this research agenda is to understand the mechanisms through which these institutional features influenced social behavior over our period and to examine how different historical factors interacted with each other. For instance, how did colonial power interact with traditional caste and social divides? Did the land tenure system introduced by the British create new class identities or did it give greater power to existing social elites? The meritocratic system of public exams introduced under colonial rule certainly created a new metropolitan-based elite. How did it influence power equations in rural areas? Finally, how have the myriad economic and social changes (land reform, industrialization, affirmative action and poverty alleviation programs) introduced after political independence affected the nature of social divisions that the country inherited? These are among the important open questions that remain.

[3]One caveat is that the instrument for the non-landlord variable is essentially just a non-linear function of the date of conquest by the British, which may also have a direct effect on economic outcomes. However, our regressions include a dummy for being conquered before 1848, which controls to some extent for direct effects of being conquered early. The instrument for the non-British variable might be invalidated if a king is more likely to die without an heir in certain types of districts. Iyer (2003) deals with this problem by controlling for deaths of kings with heirs in the same period, as well as for deaths of kings without heirs in other periods, and shows that the results are robust to these. Here we stick to the simpler specification.

References

Alesina, Alberto, Reza Baqir, and William Easterly (1999) 'Public goods and ethnic divisions.' *Quarterly Journal of Economics* 114(4), 1243–1284

Banerjee, Abhijit, and Lakshmi Iyer (2003) 'History, institutions and economic performance: The legacy of colonial land tenure systems in India.' *mimeo*

Banerjee, Abhijit, and Rohini Somanathan (2001) 'Caste, community and collective action: the political economy of public good provision in India.' *mimeo*

—— (2004) 'The political economy of public goods: Some evidence from India.' *mimeo*

Bardhan, Pranab, and Jeff Dayton-Johnson (2002) 'Unequal irrigators: Heterogeneity and commons management in large-scale multivariate research.' In *The drama of the commons,* ed. Elinor Ostrom et al (Washington DC: National Academy Press) pp. 87–112

Iyer, Lakshmi (2003) 'The long-term impact of colonial rule: Evidence from India.' *mimeo*

Khwaja, Asim (2002) 'Can good projects succeed in bad communities? Collective action in public good provision.' *mimeo*

Miguel, Edward, and Mary Kay Gugerty (2004) 'Ethnic diversity, social sanctions, and public goods in Kenya.' *Journal of Public Economics, forthcoming*

TABLE 1: IMPACT OF HISTORICAL FACTORS ON AVAILABILITY OF PUBLIC GOODS

		OLS coefficients			IV coefficients		
	Mean	Non-British	Non-landlord	Caste	Non-British	Non-landlord	Caste
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Education							
Primary school	0.794	-0.007	0.030	-0.199***	-0.005	0.036	-0.194**
		(0.021)	(0.020)	(0.071)	(0.038)	(0.038)	(0.078)
Middle school	0.254	0.040**	0.043**	-0.187*	0.019	-0.023	-0.244**
		(0.019)	(0.018)	(0.109)	(0.037)	(0.041)	(0.108)
High school	0.126	0.023**	0.045***	-0.185***	0.025	0.034*	-0.194***
		(0.009)	(0.010)	(0.058)	(0.016)	(0.020)	(0.058)
Adult literacy center	0.083	0.062***	0.015	0.145**	0.080***	-0.022	0.112
		(0.019)	(0.014)	(0.071)	(0.030)	(0.026)	(0.071)
Communication							
Post office	0.301	0.112***	0.087***	0.093	0.127*	0.028	0.043
		(0.027)	(0.020)	(0.151)	(0.066)	(0.050)	(0.144)
Telegraph office	0.051	0.019***	0.019***	0.014	0.023**	0.003	-0.000
		(0.006)	(0.007)	(0.036)	(0.011)	(0.014)	(0.035)
Phone connection	0.102	0.062***	0.058***	0.048	0.089***	0.053	0.042
		(0.011)	(0.015)	(0.081)	(0.022)	(0.032)	(0.074)
Transport							
Bus service	0.440	0.172***	0.226***	-0.466***	0.236***	0.156***	-0.528***
		(0.029)	(0.026)	(0.109)	(0.055)	(0.049)	(0.115)
Rail service	0.020	-0.000	0.002	-0.031**	-0.001	0.003	-0.030**
		(0.002)	(0.002)	(0.013)	(0.004)	(0.005)	(0.013)
Paved road	0.461	0.034	0.155***	-0.193*	0.099**	0.097**	-0.245**
		(0.028)	(0.025)	(0.112)	(0.050)	(0.048)	(0.123)
Health							
Primary health subcenter	0.096	-0.041**	0.015	-0.135*	-0.014	-0.015	-0.161**
		(0.016)	(0.012)	(0.075)	(0.026)	(0.022)	(0.078)
Primary health center	0.052	0.014***	0.022***	0.007	0.011*	-0.007	-0.018
		(0.004)	(0.006)	(0.029)	(0.006)	(0.009)	(0.024)
Dispensary	0.060	0.017***	0.011*	0.018	0.007	0.010	0.018
		(0.006)	(0.006)	(0.035)	(0.009)	(0.015)	(0.035)
Hospital	0.029	0.003	-0.006*	-0.024	0.008	0.003	-0.016
		(0.003)	(0.004)	(0.023)	(0.006)	(0.008)	(0.022)
Nursing Home	0.021	0.002	-0.007	-0.010	-0.010*	-0.005	-0.008
		(0.005)	(0.005)	(0.032)	(0.006)	(0.009)	(0.029)
Maternity and child welfare center	0.033	-0.016**	0.013**	0.021	0.002	-0.003	0.007
		(0.008)	(0.006)	(0.027)	(0.009)	(0.009)	(0.026)
Child welfare center	0.022	-0.003	0.005	-0.021	0.007	-0.002	-0.027
		(0.007)	(0.005)	(0.027)	(0.007)	(0.008)	(0.032)
Family planning center	0.031	0.005	0.011*	0.021	0.009	-0.015	-0.002
		(0.006)	(0.006)	(0.032)	(0.008)	(0.010)	(0.031)

TABLE 1 (CONTINUED)

		OLS coefficients			IV coefficients		
	Mean	Non-British	Non-landlord	Caste	Non-British	Non-landlord	Caste
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Water							
Well	0.724	-0.218***	-0.169***	0.701***	-0.182***	-0.208***	0.667***
		(0.038)	(0.036)	(0.178)	(0.064)	(0.073)	(0.187)
Handpump	0.597	-0.182***	-0.171***	0.812***	-0.208**	-0.531***	0.504**
		(0.051)	(0.049)	(0.233)	(0.094)	(0.083)	(0.253)
Tubewell	0.230	0.074*	0.007	-0.103	0.037	0.258***	0.115
		(0.039)	(0.044)	(0.178)	(0.074)	(0.094)	(0.197)
Tap	0.244	0.070**	0.167***	-0.915***	0.042	0.148**	-0.931***
		(0.032)	(0.028)	(0.158)	(0.060)	(0.062)	(0.160)
Tank	0.151	0.021	-0.056	0.403***	-0.007	-0.035	0.421***
		(0.034)	(0.034)	(0.119)	(0.063)	(0.061)	(0.125)
Power							
Power for domestic use	0.717	0.116***	0.240***	-0.330**	0.126	0.252***	-0.320**
		(0.042)	(0.031)	(0.131)	(0.083)	(0.061)	(0.140)
Power for agricultural use	0.552	0.181***	0.175***	-0.008	0.286***	0.082	-0.091
		(0.051)	(0.039)	(0.159)	(0.095)	(0.081)	(0.177)
Power for all uses	0.338	0.223***	0.160***	-0.350**	0.273***	0.142**	-0.367**
		(0.038)	(0.033)	(0.150)	(0.070)	(0.071)	(0.159)
Number of districts	391	391	391	391	391	391	391

Robust standard errors in parantheses. ***, **, * indicate significance at 1%, 5% and 10% respectively.
The dependent variable for all regressions is the proportion of villages in the district which have the public goods listed above. Column (1) shows the mean proportion of villages having the public good.
"Non-British" is a dummy which equals one if the district was under indirect colonial rule.
"Non-landlord" refers to a variable which measures the proportion of the district where land revenue was not under the control of landlords in the colonial period. This variable is set to zero for non-British areas.
"Caste" measures the extent of social diversity (caste and religious fractionalization index) in the district.
All regressions control for geographic variables (maximum and minimum temperatures, average annual rainfall, a dummy for coastal districts, the proportion of wasteland area that is mountainous, barren or rocky, and sandy); demographic variables (average village population, number of villages in the district, proportions of Muslims, Christians, Sikhs, Scheduled Castes, Scheduled Tribes and Brahmans); the gini coefficient of land ownership; and a dummy for being annexed by the British before 1848.
The instrument for non-British dummy in the IV regression is a dummy for whether the ruler died without a natural heir in 1848-56. The instrument for non-landlord proportion is a dummy for whether the British took control of land revenue between 1820 and 1856.