The Allocation of Publicly- Provided Goods to Rural Households in India: Persistent Effects between 1981 and 1991.

Roger Betancourt and Suzanne Gleason

October 2004

Introduction.

In this brief paper we update the results of our earlier work on the allocation of health and education inputs in India's rural areas. The earlier results relied on Indian 1981 census data and on a novel political data set. The new results rely on using 1991 census data now available as well as on the earlier political data set, which includes the relevant election results for the more recent period.

Several issues make this task worthwhile. First, the general topic continues to attract interest in the profession. Indeed, the World Bank devoted its most recent World Development Report (2004) to the general topic of "Making Services Work for Poor People". As part of its approach to the topic, it identifies a range of failures of service delivery. One of the failures it highlights is services that are not responsive to clients for a variety of reasons. Among the reasons it mentions is heterogeneity among clients that affect the relation between clients and providers. The report cites (p. 25) as one of the pieces of evidence in the case of India our earlier result that districts with lower proportions of scheduled castes and religious groups (Muslims) have fewer doctors and nurses per capita. Thus, the first substantive topic to be addressed is the extent to which this result continues to hold in the subsequent decade.

India's institutional structure assigns a critical role to the states in allocation decisions with respect to these health and education inputs. Outcomes with respect to these inputs as well as with respect to health and education status are known to differ dramatically across some Indian states, for example Uttar Pradesh and Kerala (Dreze and Gazdar, 1996; Dreze and Sen, 2002). Our earlier results with the 1981 Census were consistent with this view in that state effects explained a substantial part of the variation in the allocation of these inputs across

districts. Thus, a second substantive issue to be addressed is the extent to which this result continues to manifest itself with the 1991 census.

It is well known in the literature that public spending on services has difficulty in reaching poor people, for example Filmer, Hammer and Pritchett (2000, 2002). One of the two weak links in the delivery of primary health care identified by these authors is the efficacy of the public sector in providing the services. In our earlier work with the 1981 Census we found that variables proxying for aspects of bureaucratic structures were important in determining allocations of health inputs to rural areas of districts. Since the available amount of inputs of essential personnel affects efficacy, a third substantive issue to be addressed here is the extent to which these bureaucratic features continue to be a factor in determining these allocations in 1991.

The 2004 WDR identifies four relationships of accountability in the delivery of public services: from politicians to citizens; from organizational providers (bureaucracies) to politicians; from bureaucrats to their bureaucracies; and from the bureaucracy to the citizen/client. In our earlier results for 1981 some of the proxies for the existence of the first accountability relationship seemed to matter in determining the allocation of health and education inputs to rural areas of districts, i.e., in affecting the fourth link. Thus, a fourth substantive issue to be addressed here is whether proxies such as the ratio of female to male voters and voter turnout continue to have the same effects with the 1991 census.

Whether or not these relationships evolve through time is a question that has attracted the attention of other researchers on India. For instance, in a clever paper Foster and Rosenzweig (2001) argue that in India during the period 1980 to 2000 there was both democratization and

decentralization in the distribution of local public goods to villages. They arrive at this conclusion by first developing a general equilibrium model in which to embed either a two party representative democracy probabilistic voting model for choosing among three public goods or an authoritarian alternative. Then, they estimate a number of empirical relationships suggested by the model which provide evidence that democratization, measured by the share of landless in a village, has an effect when interacted with the existence of decentralized local institutions on the level of public goods -- namely roads-- favored by the landless during this 20 year period. In a much simpler setting, we'll examine the extent to which the cross-section relationships for 1981 continue to manifest themselves in 1991.

By addressing the latter issue we are led into a methodological discussion of the general reliability of cross-section results. Suffice it to say here that cross-section results are viewed with skepticism because of their inability to control for heterogeneity and its possible correlation with unobservable variables and for the endogeneity of some variables. Having two cross-sections ten years apart substantially mitigates these problems if the same relationship appears to hold in both decades. More generally, our results are to be taken in the same spirit as cross-country regressions but at a lower level of aggregation. In this connection, it is worthwhile to note Pritchett's (2000) discussion of what can be learnt from cross-country regressions and the dangers of applying panel data methods in these settings.

In the next section, we discuss the comparability of the two censuses. Table 1 shows the similarities and differences in the descriptive statistics associated with the variables for the

comparable set of 306 districts in both decades.[1] There has been a decrease in the mean number of doctors per 10 persons available to rural areas of these districts and an increase in the mean number of teachers and nurses. Both voter turnout and the ratio of female to male voter turnout exhibit increases in the mean and decreases in the standard deviation between 1981 and 1991. The proportion of scheduled castes and Muslims in rural areas of these districts increases slightly during the decade. The proportion landless also increases as do rural land per person in the rural areas and the proportion urban. This last observation suggests that there was migration out of the rural areas of these districts and that it was the better placed cultivators rather than the farm laborers who migrated in greater proportions.[2]

In Section 2, we discuss the similarities and differences in the results for our sample of 306 comparable districts. Section 3 presents an alternative specification of the relationship between the inputs allocated to health an education and the explanatory variables. This alternative provides a test of the robustness of our results over both decades. A concluding remarks section briefly highlights our main results and their implications.

1. Comparability Replications

In this section we revisit the results obtained in our earlier work using the 1981 data to put them on a comparable basis with the results that emerge from the 1991 data. Our first comparability problem was that the data we had on agricultural productivity for the 1981 data

---

[1] In an appendix we provide a succinct description of our 1981 specification. For a more extended discussion of both the institutional structure and the rationale underlying the inclusion of variables the reader is referred to our earlier work.

[2] While rural urban migration normally takes place from lower income areas to higher income ones, the migrants from a particular rural region are often wealthier than the ones who stay behind, Bausell (1975).

came from a different source than the census data on districts and is not available for 1991. Our approach to this issue was to repeat the estimations in our 1981 study dropping the agricultural productivity variables to see if this would affect our main results. The outcome of this experiment is presented in the first two columns of Table 2, where we present the results for the final model selected by the test of over-identifying restrictions in our earlier study in column 1 and the new results in column 2. Interestingly, the same test suggested the same model once agricultural productivity is dropped from the analysis: namely, one where rural land per person squared was included as having a direct impact on outcomes for doctors and nurses and the proportion of rural Muslims squared was included as having a direct impact for teachers.[3]

In all three cases, the results are basically the same. For instance, the included coefficients that are statistically significant (t-ratios > 2) when agricultural productivity is included are also statistically significant when it is excluded. Not only are the signs of these coefficients the same with and without agricultural productivity but the coefficients differ only in or after the first decimal. Not surprisingly, the main consequence of dropping this variable is in increasing the proportion of the explanation that can be attributed to the state dummies, especially in the equations for doctors and teachers. Of course, there is a bit more variability in the results for the statistically insignificant variables.[4]

What are the implications for our substantive results from the earlier study? We had stressed the importance and robustness of state effects and this remains the same. We had

---

[3] We used the overidentification test suggested by Davidson and Mackinnon (1993, pp. 232-237).

[4] For instance, the greatest difference is a change in the sign of the point estimate. This happens twice and both times it occurs in the teachers' equation.

stressed the importance and robustness of the effects of caste and religion and that remains the same for religion but the new results raise doubts about the role of caste in the case of teachers. The preferred specification for teachers had showed signs of multicolinearity in the earlier study and this had led us to minimize the importance of the result for caste in this specification. The new results don't show signs of multicolinearity but the result for caste in this equation continues to be statistically insignificant, suggesting that the effect of caste is less robust than the effect of religion as it does not hold up in all specifications in the teachers' equation. The effect of the variables that proxy for bureaucratic features are equally important with or without agricultural productivity included and the same is the case for the political variables. The results for the proportion landless are also the same. In sum, the overall relationship and the four substantive results are quite similar with or without agricultural productivity.

Our second comparability problem arises with respect to the number of districts used in the analysis. In the 1981 Census, Jammu and Kashmir districts were included but the 1991 Census skipped Jammu and Kashmir due to civil unrest. In addition there were 6 districts in 1981 that could not be reasonably recreated from the 1991 districts. Both these problems reduced our sample from 325 districts to 306 districts.[5] Our approach to this issue was to repeat the estimation of our earlier results without agricultural productivity using 1981 data with only the 306 districts for which we could compare 1981 with 1991. The results are presented in column 3 of Table 2.

A comparison of the results in columns 2 and 3 provides an indication of the sensitivity

[5] The first problem also eliminates one of our states, Jammu and Kashmir, from the set of state dummies.

of our earlier results to the change in the number of districts. A major difference arises because the model selected by the test of over-identifying restrictions is different between columns 2 and 3. That is, in column 3 (with the reduced number of districts) the test leads to the choice of a specification where rural land per person squared has no direct effect on doctors or nurses available and the proportion of Muslims squared has no direct effect on teachers. They only affect the results as part of the set of instruments for the two political variables viewed as endogenous because of the stock nature of the dependent variables.[6] Thus, the first difference we find is that nonlinearities play no direct role in the selected specification with the reduced number of districts. This holds for all three equations.

Not surprisingly in light of the previous difference in results, there are other important differences in the statistical results for all three equations with the reduced number of districts. First in every equation there is at least one variable that turns from statistical significance to statistical insignificance using a t-ratio of 2. In the doctors' equation, this is the case for the caste variable; in the nurses equation, this is the case for the caste variable as well as for the number of voting constituencies and rural land per person; finally, in the teachers' equation, this is the case for proportion landless and rural land per person. Second, in the nurses equation there is one variable (voter turnout) that changes sign while remaining statistically significant in both cases. Finally, there are other variables that change sign in the nurses' and teachers' equation but they remain statistically insignificant (t-ratio less than 2) in both cases.

---

[6] Voter turnout and the ratio of female to male voter turnout are measured for the 1977-78 elections but in a cross section precedence in time does not ensure exogeneity when there are persistence effects, Deaton (1997). Since our dependent variables are stocks, persistence effects are likely.

What happens to the substantive implications of the earlier statistical results? Fortunately, these implications survive the reduction in the number of districts but some require modifications. The first major implication that continues to hold is that state effects are extremely important in explaining variations in the allocation of doctors, nurses and teachers to the rural areas of districts of India. While the relative variation explained without state effects increases for doctors and teachers, it decreases for nurses. In every equation, however, it is less than 50% of the variation explained. Thus, state effects explain more than 50% of the variation. Finally, in all three cases the total variation explained is greater with the reduced number of districts.

The second major implication that continues to hold is that rural areas of districts with a greater proportion of Muslims receive smaller allocations of doctors, teachers and nurses. The other half of this proposition, however, is not robust. We can't say the same thing about scheduled castes. For, while the sign of the result remains the same, this result is now statistically insignificant in all three equations. Just as before, the third substantive result holds in that variables capturing bureaucratic rules matter but not in all equations. The proportion of the district that is urbanized continues to increase the number of doctor and nurses allocated to rural areas of the district and to have no effect on the number of teachers. In contrast to the earlier results, however, rural land per person matters in the doctors equation but no longer in the other two equations.

In regard to the political variables, the effect is the same as before in that 1) increases in the ratio of female to male voters increases inputs of doctors and teachers while it has no effects on nurses and 2) increases in voter turnout decrease inputs of teachers and have no effect on

doctors. Yet, it is different from before in that voter turnout changes sign in the equation for nurses while remaining statistically significant in both cases. The number of voting constituencies in a district, which mattered for nurses, no longer maters in any of the three equations. Finally, the proportion landless, which mattered for teachers but not for doctors and nurses, no longer matters in any of the three equations.

2. Similarities and Differences in the Relationship Over the Decade.

We are now in a position to investigate the stability of our relationship between 1981 and 1991. We estimated exactly the same model as before, equation (A1) in the Appendix, using the 1991 data for the 306 districts. The results are presented in column 4 of Table 2. A comparison of columns 3 and 4 allows us to ascertain the extent to which the results from the 1981 analysis persist or are robust to the changes that occurred during the decade. Interestingly enough, the specification selected by the data through the test of over-identifying restrictions is the same in 1981 as in 1991. That is, the selected model is the one where there are no nonlinearities that have a direct effect on the dependent variables and all six instruments are valid. Thus, regardless of the effects of omitted variables, measurement errors or unobserved heterogeneity, the overall statistical relationship in the cross-section seems to be the same in both decades.

With respect to the substantive results, the first similarity that persists between 1981 and 1991 is the importance of state effects in explaining the variation in the number of doctors, nurses and teachers allocated to rural areas of these districts. In all three equations state effects account for 28% or more of the explained variation in these allocations. The relative proportions explained by other factors, however , change dramatically during the decade. In the case of

doctors there is a substantial decrease so that state effects explain most of the variation while in the case of nurses and teachers there is a substantial increase so that factors other than state effects account for 55% and 72% of the explained variation, respectively. Finally, the overall variation explained by the model increases for nurses and decreases for doctors and teachers.

The second substantive similarity that persists between 1981 and 1991 is the robustness of religion in terms of the ability of the proportion of Muslims residing in a district to explain decreases in the numbers of doctors, nurses and teachers allocated to rural areas of these districts and the lack of robustness of scheduled castes in explaining these allocations. The third substantive result on the role of bureaucratic rules becomes weaker. Just as before rural land per person continues to matter for the allocation of doctors but not for that of nurses and teachers. On the other hand, the positive effect of the proportion of a district that is urbanized does not persist between 1981 and 1991. This effect disappears from both the doctors' and the nurses' equation and it continues to not matter in the teachers' equation. A fourth substantive result that persists across the decade is that the effects of the political variables vary in robustness and across equations. One persistent result is the positive effect of the ratio of female to male voters on the allocation of doctors and teachers to rural areas of a district and its lack of effect on nurses. Another one is that voter turnout decreases the number of teachers allocated to rural areas of a district in both decades. Its effect on doctors is statistically insignificant in both periods and its effect on nurses becomes statistically insignificant in 1991. The number of voting constituencies does not matter in either decade. Finally, a substantive result that does not persist across the two decades is that the proportion landless, which failed to matter in 1981, becomes a positive factor in the allocation of doctors in 1991.

What do we learn from this comparison across the decades? First, given our selection of functional form the same overall statistical relationship persists across the two decades. This is not a trivial finding. For instance, as we saw in the previous section, a small change in the number of districts in 1981 was sufficient to invalidate this result in the cross-section relation for 1981. Similarly, state effects and religion were important determinants in the allocation of these inputs for all inputs and in both decades. This result provides guidance for future research and also for policy. With respect to future research it suggests that we learn more about state effects and what they mean. With respect to policy it shows more convincingly how religious discrimination hurts the accessibility to basic health and education inputs of poor households.

The political variables had persistent effects but not for all inputs. The ratio of female to male voter turnout affected positively the availability of doctors and teachers in both decades. Voter turnout itself affected negatively the availability of teachers in both decades, which is on the surface a peculiar result.[7] The main persistent effect of bureaucratic variables is that rural land per person affected positively the availability of doctors in both decades. Overall these results suggest that a principal agent view of the relation between the electorate and state decision makers and between state decision makers and their bureaucratic agents is a useful way to think about these allocation decisions in both decades.

What about the non-persistent results?[8] They can be the result of legitimate changes in

---

[7] One explanation maybe that the increase in voter turnout leads to increased expenditures in areas such as infrastructure and, thus, requires adjustment in social expenditures categories. Teachers are ten times more abundant than doctors or nurses in both decades (see Table 1); hence, it may be easier to decrease their numbers.

[8] In some sense coefficient estimates that are not significantly different from zero in both decades could be viewed as persistent results. On the other hand, since these coefficient

the relationship across the decade or they can be the result of spurious relationships due to the problems of estimation that arise in a cross-section.[9] Most non-persistent results involved a statistically significant variable in 1981 becoming insignificant in 1991. The two exceptions are proportion landless in the doctors' equation and proportion urban in the teachers' equation, which turn from not statistically significant in 1981 to significant in 1991. Since the results hold for only one of the two decades, they should be viewed with some skepticism. Moreover, the results that go from statistically significant in 1981 to insignificant in 1991 are less interesting on two counts: they may be spurious or they may no longer be relevant.

With respect to the two exceptions, the possibility that they indicate a new aspect of the relationship can not be discarded. Note that the proportion landless result is consistent with the Foster and Rosenzweig (2001) claim of increasing democratic decentralization, if one thinks that availability of doctors is one of the public goods the landless would demand. It is also consistent with the view in our earlier paper that it is easier for the landless to organize as a homogeneous group and exert pressure to increase the number of doctors. The result on the proportion urban is an attractive candidate as a spurious result, since its sign contradicts our earlier interpretation as an indicator of a lower cost of acquiring the input.

---

estimates have wider confidence intervals than the other ones, we don't have a great deal of confidence in the stability of these results and, thus, we prefer to ignore them in this section's summary remarks. Perhaps they should be categorized as persistent 'non-results'.

[9] Incidentally, one could also argue about the persistent results that they are spurious and don't represent a lack of change in the relationship, but the argument is considerably weaker . It requires the omitted variable, the measurement error or the unobserved heterogeneity to retain the same cross effects with the included variables across the decades and to do so in such a way that statistical significance is preserved.

3. An Alternative Specification

One way of looking at the robustness of our results is to specify the model in a somewhat different way and observe which results continue to hold. In equation (A1) we included the error term by adding it to the logistic specification of the model. Alternatively, the same deterministic model could have been specified as

$$y_{ij} = [e^{\beta X(i,j)} /(1 + e^{\beta X(i,j)})] \qquad (1).$$

We could have then transformed the equation as follows

$$\ln [y_{ij} / (1- y_{ij})] = \beta X_{ij} \qquad (2)$$

and engage in linear estimation by adding the error term at this stage to obtain

$$\ln [y_{ij} / (1- y_{ij})] = \beta X_{ij} + u_{ij} \qquad (3)$$

Models (A1) and (3) are not the same, of course, but they are similar. There are two main differences between (A1) and (3). First and foremost, in (A1) the dependent variable at time t, $y_{ij}(t) = [n_{ij}(t) /pop_i(t)]$, is the ratio of the number of a type of service providers in a district , $n_{ij}(t)$, to the number of persons in that district, $pop_i(t)$. By contrast in (3) the dependent variable at time t is a function of $[y_{ij} / (1- y_{ij})] = n_{ij}(t)/[pop_i(t) - n_{ij}(t)]$, which is the ratio of the number of a type of service provider in a district to the number of persons in the district that are not providers of that type of service. Second, in (3) one tries to explain the logarithm of this ratio whereas in (A1) one tries to explain a (slightly different) ratio itself. The advantage of this specification is that the type of service provider appears only in the numerator of the ratio and that the estimation procedure can be a linear one.

We estimated the alternative specification (3) with our data by 3SLS for both 1981 and 1991. The results are presented in Table 3, together with the results from forcing the relationship

to be the same in 1981 as in 1991 (labeled Pooled). In comparing the two sets of estimates our measuring rod was a focus on whether or not a comparison of the alternative specifications would lead to different conclusions about the results than those presented in the previous section.

State effects in the alternative specification continue to be persistent and important in all three equations. The relative importance, of course, does not remain the same. For instance, state effects in the alternative specification for teachers explain over 50% of the 1991 variation whereas these effects explain 28% of the variation in the original one for 1991. Similarly, the effect of religion continues to persist in both specifications for all three equations. The effect of one of the political variables, the female to male voter turnout, is the same with both specifications: It has a persistent positive effect in the doctors' and teachers' equation. The same is true for one of the bureaucratic variables, rural land per person, which continues to have a persistent effect in the doctors' equation.

Ironically, voter turnout is the source of minor differences in results. In the teachers' equation, it is one persistent result that fails to hold in the alternative specification when one uses a 5% or lower level of significance.[10] In the nurses' equation it is one result that is persistent in the alternative specification at the 5 % or lower level of significance.[11] Most results that were statistically insignificant in both decades with the original specification continue to be with the alternative one. The three exceptions are scheduled castes, which becomes significant in the equation for doctors in the alternative specification for 1981 but not in the one for 1991; proportion landless, which exhibits the same behavior as scheduled castes but in the equation for

[10] It would hold if one were to use a 10% level.

[11] It would hold in the original model also if one were to use a 10 % level of significance.

nurses; and rural land per person in the teachers' equation, which becomes a non-persistent result in the alternative specification. Non-persistent results in the original specification remain pretty much the same in the alternative one.

A comparison of the results for each decade with the pooled results reveals a substantial loss in explanatory power as a result of pooling in all three equations. State effects explain a greater proportion of the variation in 1991 for doctors and nurses than in the pooled equation and they explain a lower proportion of the variation in 1991 for the teacher's equation than the pooled equation. By contrast, the other explanatory variables explain a greater proportion of the variation in each decade than in the pooled equation for doctors and teachers and for nurses in 1981 but not in 1991.

Concluding Remarks

What have we learnt from this exercise? First, state effects are persistent and important in determining the allocation of these inputs to districts, but they do change over time. Hence, as in the case of Pritchett's (2000) analysis of growth regressions, these results suggest that in India analyses of historical episodes in the evolution of individual states policies with respect to the assignment of these inputs to districts is a valuable research activity. For instance, Kutty (2000) provides an example of insightful analysis with respect to health policy in Kerala by a non-economist. It documents the invaluable role of the public sector in providing health service inputs up to the mid-1980's and its inability to do so since that time.

Second, in making services work for the poor in India, heterogeneity with respect to religion, and more specifically with respect to the Muslim religion, is far more powerful and

persistent in making services non-responsive to poor clients than other forms of heterogeneity, including heterogeneity with respect to scheduled castes. Third, a bureaucratic rule based on geography is a persistent determinant of the allocation of doctors to less densely populated rural areas of districts in India. The application of a similar rule to other inputs may increase the ability of public spending to reach the poor in remote areas.

Finally, the accountability of politicians to female clients has a persistent effect increasing the provision by India's state bureaucracies of doctors and teachers to the rural areas of districts where these clients reside. In contrast, the main persistent effect of accountability to voters in general is that India's state bureaucracies decrease the number of teachers and perhaps nurses allocated to rural areas of district where these voters reside when voter turnout increases. This suggests that government provided teachers in rural areas may be viewed as inferior public goods. Incidentally, the net effect of an increase in female voter turnout, given male voter turnout, is an increase in the number of doctors and a decrease in the number of teachers. Hence, our results are consistent with the results of Chattopadhyay and Duflo (2004) who found that health professionals were monitored more closely and teachers less closely in areas where women had greater influence than in areas where women had less influence. In their work influence was determined by whether or not positions as chief of the village councils were reserved for women.

Appendix: The Original Specification and the New Data Set .

In our earlier work, we estimated the following model by NL3SL

$$y_{ij} = e^{\beta X(i,j)} /(1 + e^{\beta X(i,j)}) + \epsilon_{ij} \quad (i = 1, 325; j = 1,2,3). \qquad (A1)$$

While j identifies the equation for the publicly provided input, i.e., doctors, nurses and teachers, i identifies the unit of observation, i.e., India's administrative districts. Thus, y_{ij} was measured as the amount of the jth of these inputs available in rural areas of a district for every 10 persons living in rural areas of this district. They can be viewed as indexes of the quantities of medical and educational services available to poor households in India's rural areas. We were able to measure these same exact variables with the data provided by the 1991 Census.

In 1981 there were 9 explanatory variables in the vector X, in addition to the intercept, and 16 dummy variables to capture state effects for the 17 states among which our districts were distributed. In 1991 there are 8 explanatory variables in the vector X, in addition to the intercept,

and 15 dummy variables to capture state effects for the 16 states among which our districts are distributed.

Of the 9 explanatory variables in the earlier study we were able to measure exactly the same variables in 8 cases. Five of the 9 came from the 1981 decennial Census, just as the dependent variables do. We were able to obtain exactly the same variables in the 1991 Census. Two of them, the proportion of schedule castes and Muslims in the rural areas of districts, aimed to capture selectivity in the allocation of these medical and educational inputs with respect to these groups. Another two of these variables, the proportion of urban population in a district and the area of rural land per person in rural areas of a district, aimed to capture bureaucratic rules or procedures based on population or geography that might determine these allocations. The last of these 5 variables is the proportion of landless (farm workers who are laborers and not cultivators) in a district. This variable and the one we could not obtain (agricultural

productivity) aimed at capturing the role of interest groups in affecting these allocations.

In our earlier study agricultural productivity information had been obtained from a data base other than the actual 1981 decennial census and this data base has not been duplicated for1991. This raises the methodological issue of how to account for the absence of this variable. This methodological issue is addressed in Section 1, by showing that the results for the remaining variables are not substantively affected as a result of dropping this one.

Since the remaining three variables come from a data set constructed by us, it was possible to construct them in a similar manner for use with the 1991 Census. Election results for each district are taken from the legislative assembly election[12] immediately preceding the 1991 census (the elections included are primarily 1987-1990[13]). Results are reported by voting constituency. Voter turnout measures were generated by summing the constituencies contained in each district, in total and by gender. Voting constituencies were matched to census districts using the correspondence published in Singh & Bose (1987-88).[14]

Finally, 'attrition' takes place among our units of observation between 1981 and 1991 for two different reasons. First, civil unrest led to the decennial census in 1991 not taking place in

---

[12]The election results from all legislative assembly elections since 1951 are published by the Election Commission of India (http://eci.gov.in/infoeci/key_stat/keystat_fs.htm).

[13]Elections results used in the original research were from the 1977/78 elections.

[14]Singh, V.B., and Shankar Bose. 1987-88. *Elections in India: Data Handbook on Lok Sabha and Vidhan Sabha elections 1952-85*. New Delhi: Sage Publications. For a detailed description of this process for the 1981 data set, please refer to Gleason 1996. We mapped the constituencies for the more recent elections to their 1981 districts in order to provide continuity over the period. The 16 states which are included in the analysis did not see any significant redrawing of constituency boundaries in the intervening years. This is not the case for all states. Subsequent elections have seen more change in constituency boundaries making the election-census mapping more difficult.

some areas. More specifically, districts in the state of Jammu and Kashmir had to be excluded from the data set. Second there were some changes in the definitions of the districts between 1981 and 1991, which meant that 6 of the 1981 districts could not be recreated in the 1991 data.[15] In Section 1 we address this comparability problem by estimating the same specification in 1981 with the original and the reduced number of districts.

[15] Recently, economists have started to focus attention on the fact that administrative and legal jurisdictions vary over time, e.g., Alessina, Baqir and Hoxby (2004)

References

Alesina, A., Baqir, R., & Hoxby, C. (2004). Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy*, 112, 384-396.

Bausell, C. (1975). *The Decision to Migrate under Uncertainty: A Case Study of Rural- Urban Migration in Tropical Africa.*. Unpublished doctoral dissertation, University of Maryland, College Park

Betancourt, R.,& Gleason, S. (2000). The Allocation of Publicly-Provided Inputs to Rural Households in India: On Some Consequences of Caste, Religion and Democracy. *World Development*, 28, 2169-2182.

Chattophadyay, R., & Dufflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, 72, 1409-1443.

Davidson, R.,& Mckinnon, J. G.(1993). *Estimation and Inference in Econometrics*. Oxford: Oxford university Press.

Deaton, A. (1997). The Analysis of Household surveys: A Microeconometric Approach to Development Policy. Baltimore: John Hopkins University Press.

Dreze, J., & Gazdar, H. (1996). Uttar Pradesh: The Burden of Inertia. In J. Dreze and A. Sen (Eds.), *Indian Development: Selected Regional Perspectives*. New York: Oxford University Press.

Dreze, J., & Sen, A. (2002). *India: Development and Participation*. Oxford: Oxford University Press.

Filmer, D., Hammer, J.,& Pritchett, L. (2000).Weak Links in the Chain: A Diagnosis of Health

Policy in Poor Countries. *World Bank Research Observer*, 15, 199-224.

Filmer, D., Hammer, J.,& Pritchett, L. (2002).Weak Links in the Chain II: A Prescription for Health Policy in Poor Countries. *World Bank Research Observer*, 17, 47-66.

Foster, A., & Rosenzweig, M. (2001). Democratization, Decentralization and the Distribution of Local Public Goods in a Rural Economy. mimeo, Brown University.

Gleason, S. (1996). Female Well-Being and Public Goods. Unpublished doctoral dissertation. University of Maryland, College Park.

Kutty, V. R. (2000). Historical Analysis of the Development of Health Care Facilities in Kerala State, India. *Health Policy and Planning*. 15, 103-109.

Pritchett, L. (2000). Understanding Patterns of Growth: Searching for Hills among Plateaus, Mountains and Plains. *World Bank Economic Review* , 14, 221-250.

Singh, V.B.,& Bose, S. (1987-88). Elections in India: Data Handbook on Lok Sabha and Vidhan Sabha Elections1952-1985. New Delhi: Sage Publications.

World Bank. (2004). Making Services Work for Poor People. *World Development Report*. New York: Oxford University Press.

Table 1: Descriptive Statistics

	1981		1991	
Variable	Mean	S.D.	Mean	S.D.
Voter Turnout	0.575	0.126	0.620	0.102
Female to Male Voter Ratio	0.746	0.172	0.862	0.119
Number of Constituencies	10.24	6.79	10.24	6.79
Proportion Scheduled Castes	0.179	0.080	0.189	0.080
Proportion Muslim	0.080	0.095	0.083	0.101
Proportion Landless	0.344	0.182	0.368	0.182
Rural Population Density	0.075	0.121	0.081	0.251
Proportion Urban	0.186	0.119	0.207	0.128
Agricultural output per person	0.221	0.142	--	--
Doctors	0.0051	0.0026	0.0049	0.0022
Nurses	0.0050	0.0029	0.0056	0.0038
Teachers	0.041	0.018	0.046	0.020
Number of Observations = 306				

Table 2: Comparability Replications and New Results: NL3SLS Estimates (t-stats)

Variable	original	Omit agpop	Small group	1991
Doctors				
Voter Turnout	1.29	1.22	0.94	0.016
	(1.60)	(1.51)	(0.76)	(0.15)
Female/Male Voter Turnout	0.71	0.87	0.68	1.28
	(3.02)	(3.39)	(2.32)	(2.08)
Number of Voting Constituencies	-0.082	-0.076	-0.081	0.073
	(1.66)	(1.52)	(1.67)	(1.93)
Prop. Rural Scheduled Castes	-0.23	-0.23	-0.22	-0.061
	(2.93)	(2.84)	(1.92)	(0.74)
Proportion Rural Muslims	-0.091	-0.087	-0.073	-0.056
	(3.10)	(2.80)	(2.85)	(2.82)
Ag. Output per person	-0.078			
	(1.07)			
Proportion Landless	-0.040	-0.074	-0.0099	0.22
	(0.32)	(0.57)	(0.08)	(2.79)
Rural Land per Person	0.065	0.072	0.037	0.019
	(2.61)	(2.79)	(2.23)	(2.78)
Proportion Urban	0.17	0.16	0.18	0.047
	(4.43)	(4.31)	(4.54)	(0.96)
Rural Land per Person Squared	-0.074	-0.019		
	(2.48)	(2.64)		
Pseudo R2	0.52	0.45	0.58	0.51
Pseudo R2 omitting state effects	0.14	0.01	0.17	0.03
Nurses				
Voter Turnout	2.37	2.14	-1.74	-2.52
	(2.58)	(2.44)	(2.33)	(1.90)
Female/Male Voter Turnout	0.29	0.39	0.057	1.41
	(1.07)	(1.42)	(0.24)	(1.49)
Number of Voting Constituencies	-0.15	-0.14	-0.067	-0.00037
	(2.26)	(2.16)	(1.28)	(0.00)
Prop. Rural Scheduled Castes	-0.50	-0.51	-0.084	-0.12

	(4.88)	(4.99)	(0.68)	(0.99)
Proportion Rural Muslims	-0.17 (4.68)	-0.16 (4.63)	-0.15 (5.27)	-0.12 (3.36)
Ag. Output per person	-0.15 (1.78)			
Proportion Landless	-0.20 (1.72)	-0.22 (1.87)	0.15 (1.97)	0.11 (1.53)
Rural Land per Person	0.096 (2.60)	0.10 (2.76)	0.030 (1.72)	0.0064 (0.55)
Proportion Urban	0.15 (3.69)	0.15 (3.63)	0.12 (2.26)	-0.054 (0.74)
Rural Land per Person Squared	-0.023 (2.35)	-0.024 (2.47)		
Pseudo R2	0.58	0.58	0.62	0.74
Pseudo R2 omitting state effects	0.40	0.30	0.31	0.44

Table 2: Comparability Replications cont.: NL3SLS Estimates (t-stats)

Teachers				
Voter Turnout	-2.20	-2.60	-1.84	-3.17
	(4.01)	(4.97)	(2.92)	(2.68)
Female/Male Voter Turnout	0.75	0.83	0.84	2.22
	(4.91)	(5.17)	(4.22)	(2.70)
Number of Voting Constituencies	-0.0065	0.0058	-0.019	0.049
	(0.16)	(0.13)	(0.52)	(0.73)
Prop. Rural Scheduled Castes	-0.0033	0.022	-0.080	-0.051
	(0.04)	(0.29)	(1.05)	(0.44)
Proportion Rural Muslims	-0.21	-0.22	-0.052	-0.046
	(5.39)	(5.42)	(2.36)	(2.41)
Ag. Output per person	-0.11			
	(1.66)			
Proportion Landless	0.17	0.17	0.13	0.0065
	(2.27)	(2.21)	(1.79)	(0.07)
Rural Land per Person	0.013	0.016	0.011	0.011
	(4.72)	(6.30)	(1.38)	(1.26)
Proportion Urban	0.00033	-0.0019	-0.014	-0.14
	(0.01)	(0.05)	(0.41)	(2.01)
Proportion Rural Muslims Squared	0.082	0.090		
	(5.34)	(5.90)		
Pseudo R2	0.55	0.49	0.59	0.43
Pseudo R2 omitting state effects	0.15	0.03	0.08	0.31
Number of observations	325	325	306	306

Table 3: Alternative Specification: 3SLS Estimates (t-stats)

Variable	Pooled	1981	1991
Doctors			
Voter Turnout	0.570	1.269	0.544
	(0.88)	(1.58)	(0.48)
Female/Male Voter Turnout	0.831	0.895	1.740
	(2.31)	(2.87)	(2.34)
Number of Voting	0.011	-0.057	0.057
Constituencies	(0.33)	(1.26)	(1.17)
Prop. Rural Scheduled Castes	-0.135	-0.215	-0.126
	(1.82)	(2.17)	(1.17)
Proportion Rural Muslims	-0.065	-0.074	-0.070
	(3.69)	(3.17)	(2.90)
Proportion Landless	0.068	-0.059	0.207
	(1.04)	(0.62)	(2.41)
Rural Land per Person	0.014	0.039	0.017
	(2.24)	(2.66)	(2.41)
Proportion Urban	0.109	0.170	0.050
	(3.41)	(4.62)	(0.86)
Pseudo R2	0.44	0.54	0.42
Pseudo R2 omitting state effects	0.07	0.21	0.02
Nurses			
Voter Turnout	-2.87	-2.103	-3.47
	(4.62)	(2.68)	(2.67)
Female/Male Voter Turnout	0.401	0.188	1.889
	(1.13)	(0.72)	(2.14)
Number of Voting	-0.045	-0.084	-0.00029
Constituencies	(0.94)	(1.81)	(0.01)
Prop. Rural Scheduled Castes	0.00108	-0.039	-0.044
	(0.01)	(0.34)	(0.38)

Proportion Rural Muslims	-0.108 (4.51)	-0.141 (5.39)	-0.099 (3.26)
Proportion Landless	0.219 (3.04)	0.187 (2.08)	0.103 (0.99)
Rural Land per Person	0.00364 (0.43)	0.022 (1.27)	0.00539 (0.47)
Proportion Urban	0.046 (1.24)	0.113 (2.15)	-0.072 (1.13)
Pseudo R2	0.59	0.61	0.68
Pseudo R2 omitting state effects	0.19	0.31	0.20

Table 3: Alternative Specification: 3SLS Estimates cont. (t-stats)

Teachers			
Voter Turnout	-1.455	-1.024	-2.893
	(3.22)	(1.71)	(2.70)
Female/Male Voter Turnout	1.062	0.933	2.116
	(4.02)	(4.19)	(2.75)
Number of Voting	-0.00098	-0.011	0.033
Constituencies	(0.04)	(0.33)	(0.66)
Prop. Rural Scheduled Castes	-0.0986	-0.132	-0.056
	(2.03)	(1.79)	(0.56)
Proportion Rural Muslims	-0.074	-0.076	-0.075
	(5.31)	(3.82)	(3.20)
Proportion Landless	0.0952	0.101	0.055
	(1.94)	(1.44)	(0.62)
Rural Land per Person	0.00442	0.023	0.012
	(1.14)	(3.22)	(1.62)
Proportion Urban	-0.0469	-0.019	-0.144
	(2.00)	(0.61)	(2.60)
Pseudo R2	0.51	0.61	0.43
Pseudo R2 omitting state effects	0.11	0.14	0.12
Number of observations	612	306	