ANJINI KOCHAR
Stanford University

Can Schooling Policies Affect Schooling Inequality? An Empirical Evaluation of School Location Policies in India

Introduction

The Government of India's schooling policies have long been guided by two primary objectives: to increase schooling attainment and to reduce schooling gaps, particularly those based on caste and gender. As in most developing economies which started with poor social infrastructure in rural areas, it was felt that the primary deterrents to schooling were the inadequate number of schools and the consequent distance between the average residence and a school. The government therefore made the provision of a school within walking distance from each rural household a priority.

In implementing this policy, scant attention was paid to the fact that targeting access to schools as a primary objective may constrain the government in addressing other critical aspects of schools, particularly those related to school quality. This is because decisions regarding the location of schools determine more than just access to schools, they combine with the residential structure of an economy to define the school community; hence, characteristics of schools are known to affect schooling attainment. For example, if residential centers are geographically dispersed and are of relatively small size, then the priority placed on access may imply a corresponding inability to efficiently choose school size: it will be determined by the size of the community. And, if the economy is characterized by a relatively high level of residential segregation, decisions regarding the location of schools across closely spaced communities will determine whether residential segregation implies a correspondingly high degree of schooling segregation. Under such conditions, policies to improve school quality cannot be discussed independently of school location policies.

The nature of residential communities in rural India makes this trade-off between access and quality likely. Rural India resides in habitations—distinct residential sub-divisions of a village—which vary in size but are, on an average, fairly small. Habitations are generally organized along caste lines, so that the rural economy is characterized by a considerable degree of caste-based segregation with the Scheduled Castes (SCs) and Scheduled Tribes (STs) households frequently residing in separate, smaller, sub-habitations of the village. The stated policy objective of providing a school within easy walking distance of each household in conjunction with the geographic distance across habitations requires the government to use the habitation as the basis for school mapping exercises and to adopt a policy that provides schools even to relatively small habitations.

In turn, the small size of the average habitation removes the possibility of optimally choosing school enrollment or size; it is dictated by the population size of the habitation. Reflecting this inefficiency, school size varies tremendously within a district, even within a village, and forges a similar inefficient variation in critical schooling inputs such as the number of teachers. If these inputs affect schooling attainment, then the benefits of access to schools will vary across habitations, depending on their size. And, a relatively large number of sub-optimal sized schools will imply a corresponding reduction in the average quality of schools in the country.

It is not just the size of the habitation within which a school is located that affects schooling attainment, the size of neighboring habitations also matters. The geographical distance across habitations implies that large villages commonly have more than one school, with each school serving a different habitation. In these villages, the government's school location policies translate the extensive residential segregation by caste in rural India into a corresponding system of *de facto* schooling segregation. This, in turn, is likely to affect schooling attainment and reinforce the caste-based divisions, which characterize rural India.

The socio-economic and demographic differences across the SC/ST habitations and those populated by the upper castes suggest that the government's policy of providing a school for each habitation implies a trade-off not just in terms of the average quality of schools but also in terms of schooling inequality. The size of the SC/ST habitations, even those large enough to justify a school, is significantly smaller than the average size of upper caste habitations. Schools located in the SC/ST habitations have correspondingly fewer teachers and hence are generally of lower quality than those located in the upper caste habitations. The schooling segregation which results when schools are located in the SC/ST habitations reinforces these quality differences. Schooling segregation implies that children of the upper castes will attend schools primarily populated by children from relatively well-off families, but the opposite is true for children who attend schools located in the SC/ST habitations. If the average “quality” of the student population matters for individual schooling attainment, as many believe it does, schooling segregation will increase the schooling of the upper caste children, while reducing that of the SC/ST. These school-policy induced differences in school quality are likely to explain a significant component of caste-based differences in schooling attainment.

The empirical analysis of this paper starts by exploring the relationships between the habitation size and the attributes of rural schools. The positive correlation between habitation size and school access suggested by India’s school location policy is borne out in the data. Additionally, and consistent with the hypothesis that school location policies forge a relationship between residential characteristics and other attributes of schools, the data reveal that large habitations also have an advantage in the availability of teachers; districts with larger habitations have a smaller proportion of schools with two or less teachers. Moreover, districts with larger habitations are also more likely to have relatively large SC/ST habitations and hence a greater probability of schools being located in SC/ST habitations. This suggests greater schooling segregation in these districts.

The paper then proceeds to explore the effect of these school attributes on primary enrollments. I show that the probability of a school being located within a child’s habitation of residence *does* increase enrollment. Extending this analysis, I find that small schools, with two or fewer teachers, are less likely to attract students. Further, the extent of caste segregation in schools, as measured by the probability of a school being located in a SC/ST habitation, also affects enrollment. As expected, these effects vary by caste. Schooling segregation *reduces* enrollments of SC/ST students, though the effect is not statistically significant at conventional levels of significance. However, it significantly increases enrollments by children of other caste households.

Taken together, the results of this paper confirm a relationship between habitation size and schooling outcomes, not just through the effect of size on school access but also because it influences attributes such as number of teachers. The size of neighboring habitations also matters. Specifically, the size of SC/ST habitations in a village dictates the extent of schooling segregation, which in turn further affects schooling attainment. This analysis thus confirms that school location policies affect critical determinants of school quality and contribute to the significant observed variation in school quality across regions. They also contribute to caste-based schooling inequalities. Because regions characterized by larger habitations have an advantage in access, but are also more likely to be characterized by caste segregation across schools, the benefits of improved access are magnified for the upper castes, but reduced for the SC/ST households. Extending the results of this paper, I quantify the extent of this increase.

This paper is related to several literatures. It is most closely related to the theoretical and empirical literature on the effect of community characteristics on schooling outcomes (Borjas, 1995; Case and Katz, 1991; Coleman, 1988; Fernandez and Rogerson, 1996; Miguel and Gugerty, 2005). As in this literature, I argue that community characteristics play an important role in shaping schooling attainment. However, the focus on school location policies suggests that habitation size may be as or more important than attributes such as community income and its ethnic composition which have been the focus of the literature to date. Because school size affects so many critical school inputs, this focus suggests that the influence of school location policies may be far-reaching: they can affect a very wide range of critical school inputs.

The primary contribution of my research to this literature is on the empirical side. The available empirical work on the topic documents the effect of community characteristics on schooling attainment, but generally does not show that these are mediated through the effect of the community on schooling inputs. Such a structural analysis has proven difficult, both because of the lack of school level data on school inputs and credible instruments to identify their effects. Using community characteristics to identify the effect of school inputs is questionable, since characteristics such as the ethnic composition of the community will likely affect schooling attainment directly, not just through school characteristics. Moreover, when a school serves several different residential communities, data is required not just on the community in which the school is located but also on all those that fall within the school's catchment area.

I use the rules that determine whether schools should be placed within a habitation in conjunction with habitation characteristics as the basis for identification of the effect of school characteristics. Doing so directly relates schooling outcomes to the combined influence of school location policies and habitation characteristics, and thus provides evidence on the role of school location policies in shaping schooling outcomes. I limit myself to an analysis of the effect of those characteristics which I can identify through this approach: access to schools, the number of teachers in the school, and the extent of caste-based school segregation.

The government's school mapping exercises are conducted at the level of the district, based on data collected through the All India Education Surveys (AIES) on habitations and schools. I use the same data, combined with household data on school enrollment from National Sample Surveys (NSSs) for the empirical analysis of this paper. The focus of this paper, which is on the consequences of school location policies, thus limits the analysis to the identification of district-level variation in school inputs. However, this approach also implies several benefits. It removes the necessity for school-level data, provided the appropriate district-level counterparts can be constructed. The very rich district-level data on the demographic characteristics of habitations and the habitation-wise availability of schooling inputs contained in the AIES facilitate such an analysis.

This is true also for estimates of the effect of school segregation. Recognizing that residential segregation will translate into school segregation only if schools are provided in the SC/ST habitations in addition to the other caste habitations, I use data on the availability of schools in the SC/ST habitations within the district to proxy the extent of school segregation in the district. Doing so yields several benefits. First, it emphasizes the role of school location policies, since the variable I consider is the availability of schools in the SC/ST communities, an outcome which is determined by these policies. Second, the aggregate approach of this paper automatically deals with the difference between the schooling and residential community which arises when a school's catchment area spans several communities. A disadvantage of this approach is that conclusions regarding the effects of school segregation rely on the existence of a correlation between the availability of schools in the SC/ST habitations and schooling segregation. While caste-based residential segregation is strongly suggestive of this correlation, I cannot confirm it without data on the student composition of individual schools.

The focus on school inputs as the pathway through which habitation characteristics affect schooling attainment yields a key insight which has not been sufficiently developed in the existing literature. Existing theories note that community effects on schooling will generate regional inequality in schooling attainment: children who reside in "good" communities will do consistently better than other children, of similar individual and household characteristics, who reside in "bad" communities, generating persistent schooling inequalities across communities. The focus on schooling inputs, in conjunction with the recognition that their effect on attainment can vary across households suggests an alternative form of inequality. Specifically, within any given residential community, the benefits of access to schools may be significantly higher for some households relative to others.

This research is also related to a relatively small literature which examines the determinants of schooling segregation. For example, Benabou (1996) argues that the positive effects of characteristics such as the average wealth of the community will cause schooling segregation, with richer households separating into exclusive schooling communities so as to increase the mean value of local income. In contrast, observed patterns of residential segregation in many developing economies, such as India, have not arisen as a consequence of community influences on schooling production, but have been a historical response to the desire for segregation in the use of public goods (Kochar et al., 2007). In this context, I show that existing patterns of residential segregation generate segregation in schools, but only because of the school location policy which provides each habitation, with a population above a defined threshold level, with a school.

Also related is an extensive empirical literature that examines the effect of access, or distance to school, on schooling outcomes, primarily enrollments. This literature generally finds that enrollment decisions show little response to the distance of the household to schools (Filmer, 2004). However, these results have been questioned on the grounds that they do not allow for the possibility that governments' decisions regarding school location are responsive to unobserved regional variables which may also directly affect schooling choices, including the returns to schooling.[1] Studies which have attempted to deal with this endogeneity, such as Foster and Rosenzweig (1996), have restricted their analysis to the effect of the availability of a school within a village, not allowing for residential segregation and the possibility that the location of the school within the village may be far more important to households. In addition to the difficulties associated with the endogeneity of school location, a major shortcoming of this literature, which it shares with the policy debate on the topic, is that it ignores the possibility that decisions regarding school location may combine with residential patterns to affect other attributes of schools known to influence schooling outcomes.

The rest of this paper is organized as follows. The second section describes India's school location policies, as well as policies relating to the hiring of teachers and decentralized planning which are relevant to this paper. The following section turns to the study region, the state of Uttar Pradesh, and uses the AIES data to discuss residential and schooling patterns in the state. The data used in this paper is described in the fourth section; the fifth section discusses the habitation size and school attributes while the next section discusses the empirical methodology used for writing this paper. The main results are presented in the seventh section. The last section concludes the paper.

1. Examples of such research include Birdsall (1985), Lillard and Willis (1994), Deolalikar (1997) and Alderman et al. (2001).

India's School Location and Other Schooling Policies

School Location Policy and School Quality

At independence, the Government of India inherited a very weak schooling system, with educational facilities available only for 40 percent of the children in the 6–11 years age group (Planning Commission, 1952). Many felt that this reflected an inadequate access to schools. Although the total number of primary schools in 1950–51 was 2,09,671, so that there was one school for approximately 212 children of that age group, this number did not adequately represent the distance of households to schools and hence the costs of schooling. At the time of the second AIES (1965), only 38 percent of rural habitations had a primary schools in them.

To ensure access, the government determined that every rural habitation would be "served" by a school, in that a primary school would be available either in the habitation or within easy walking distance.[2] This led to a rapid growth in the number of schools. By the Third Survey (1973), the number of habitations with primary-level schools in them increased to 44.3 percent while 75.6 percent were served by schools with a primary-level section either within the habitation or within a walking distance of 1 km. By 2002, 54 percent of rural habitations had a school, up from 50 percent in 1993. However, there remains considerable variation in access to schools across states. In Uttar Pradesh, for example, the percentage of habitations with a school located in the habitation was only 41 percent in 2002 and was as low as 29 percent in 1993.

The implementation of this policy required data on the number of habitations in rural India and the percentage of these habitations that were already served by schools. To provide this, the government initiated the AIESs with the specific objective of collecting the data necessary for school construction policies. These surveys constitute the only datasets that use the habitation as the survey unit, and which provide information on the availability of schools and schooling inputs at the level of the habitation. The data collected in any particular survey provides the basis for the government's school construction program in the period between that survey and the next survey. Thus, for example, information contained in the fifth AIES conducted in 1986 provided the basis for the government's school construction program between 1986 and 1993, as well as the data that guided the allocation of teachers in this period.

2. For the first two surveys, this meant that a school should be available within 1 mile. As of the third survey, this was changed to 1 km.

The first AIES report, submitted in 1957, specified the guidelines for construction of a school, guidelines which maintain even today. Every habitation with a population of 500 or more should have a primary school located in the habitation. For habitations with a population ranging from 300 to 499, this population criterion was combined with a distance criterion: schools should be located in habitations of this size range only if there were no primary schools (existing or proposed) within a walking distance of half a mile.

In determining the policy for the location of schools, little thought was given to the fact that decisions regarding school location would also affect other characteristics of schools, notably the school size and the socio-ethnic-economic composition of the student population. Instead, the government made clear that it considered these decisions to be separable. The second AIES survey states that its objective was to develop an approach for the location of schools, not to deal with issues such as the number of divisions or classes available or necessary in any given standard, or the optimum size of the school or class. It argued that these decisions came under the jurisdiction of the state level educational administration, in contrast to decisions regarding school availability which were under the authority of the Central Government.

By 1986, the number of primary schools had increased from 2,09,671 to 5,29,392. However, despite the significant expenditure on school construction, enrollments remained low, particularly for children from the SC/ST households. In rural Uttar Pradesh, for example, the 1999 NSS survey data reveal that of the SC/ST children between the ages of six and eleven, only 48 percent were currently enrolled in primary school. In contrast, this percentage was 56 percent for children from other castes.

Continued low schooling attainment, particularly by the SCs and STs, could reflect the very poor physical conditions of government schools in rural India. As documented later in this paper, the government's decision to separate school access decisions from concerns over optimal school size resulted in the schooling system being characterized by a large number of very small schools. In 2002, total enrollment in the average rural government school was only 114.[3] Such a small school population makes it difficult to justify the fixed costs required for investments in physical infrastructure and perhaps explains the lack of basic facilities in many of India's rural schools. For example, in 2002, 45 percent of rural government schools lacked usable

3. These data are from the Government of India's Seventh All India Education Survey.

playground facilities and only 24 percent had toilets. Further, 34 percent of sections in primary schools had no furniture for children.

Teachers

Attention to the availability of teachers and to the teacher-pupil ratio (TPR) started around the 1980s. The Steering Group on Education, Culture and Sports for the Seventh Five Year Plan (1985–90) targeted a TPR of 1:40. The availability of teachers became a central focus with the 1986 National Policy on Education. This policy introduced a scheme entitled Operation Blackboard to ensure the minimum essential facilities to all primary schools.

The introduction of Operation Blackboard in 1987–88 modified the rules for the allocation of teachers. Rather than be guided strictly on the basis of a targeted TPR, the scheme specified a minimum number of teachers who should be provided in each primary school. The original scheme called for a minimum of two teachers. An extension of the scheme in the Eighth Plan called for a minimum of three teachers in all schools with an enrollment of eighty or more. This target was subsequently relaxed so as to ensure three teachers for all schools whose enrollment exceeded 100 students.

Decentralized Planning

The Government of India has always stressed the need for decentralized planning in the context of schooling. In the early years of planning, however, this policy was implemented at the level of the state. The Sixth and Seventh Plans, for example, introduced state-specific targeting of central grants for schooling (under Centrally-sponsored schemes), making funds available on a priority basis to "educationally backward" states.[4]

The Eighth Plan, however, and the approach papers to the Plan, argued that inter-district variation in schooling was more pronounced than inter-state variation. The Working Group on Childhood and Elementary Education for the Eighth Plan therefore undertook a ranking of districts in terms of educational achievements. Arguing that there was no guarantee that funds targeted for educationally backward states would reach the backward districts within the state, it required all funds provided from central schemes to be targeted to districts (regardless of the state in which they were located) on

4. Andhra Pradesh, Assam, Bihar, Jammu and Kashmir, Madhya Pradesh, Orissa, Rajasthan, Uttar Pradesh, West Bengal, and Arunachal Pradesh belong to this category.

the basis of this ranking, and in proportion to the degree of backwardness. The rank was based on a composite index, which gave equal weight to the following four parameters (based on data available in 1991): the literacy rate, the female literacy rate, the gross enrollment ratio for primary level, and this same ratio for females. This same index guided the allocation of funds through the Eighth and Ninth Plans.

Residential Patterns in Uttar Pradesh

In 2002, there were 96,014 villages in Uttar Pradesh and more than double that number of habitations (2,01,606). Thus, on an average, each village had at least two habitations. Of the total number of habitations, 47,946 were classified as SC or ST habitations, so that SC/ST habitations accounted for 23.8 percent of total habitations. This percentage mirrored the percentage of the SCs and the STs in the state's population (23.5 percent). The total population residing in the SC/ST habitations was 3,00,06,468. This population total amounts to 89 percent of the estimated population of the SCs and the STs in the state[5] and suggests a very high degree of residential segregation by caste.

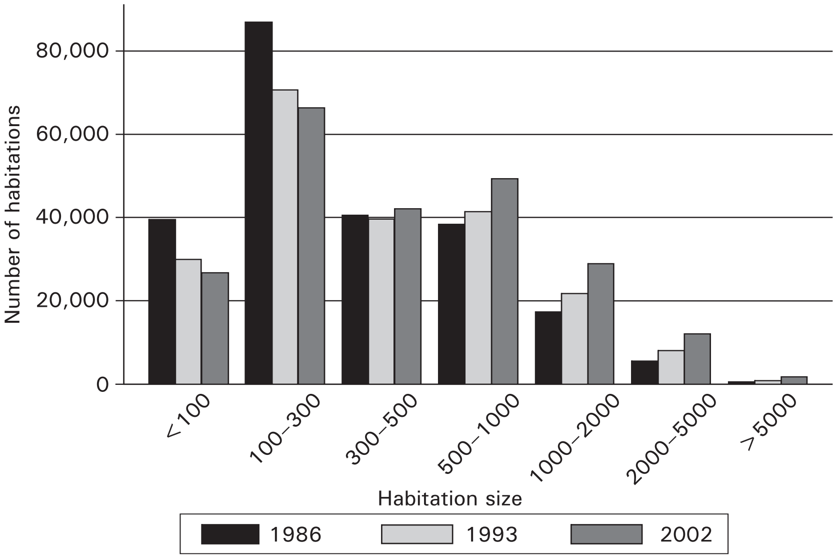
Habitations are, in general, small. Figure 1 plots the distribution of habitations by size group for the years from 1986 to 2002. The data reveal that, in 1993, 47 percent of habitations had a total population size of less than 300 while as many as 66 percent were less than 500 in size. The SC/ST habitations are smaller: the percentage of SC/ST habitations with a population size of less than 300 is as high as 56.3 percent (figure 2). Thus, the average SC or ST child is likely to reside in a habitation of smaller size than a child from a general caste.

The policy of defining a school's catchment area by the habitation of its location, in conjunction with the small average size of habitations, implies that school size is also typically very small (figure 3). In 1993, the average school size in rural Uttar Pradesh was 138 students, with 37 percent of schools having student strength of less than 100.[6] Because teacher allocations are determined by the total enrollment in the school, the small size of schools implies a correspondingly small number of teachers in each school (figure 4). The vast majority of schools in rural Uttar Pradesh have two or fewer

5. Of course, the SC/ST habitations will also include members of Other Backward Castes (OBC). Data on the exact caste composition of each habitation is not available.

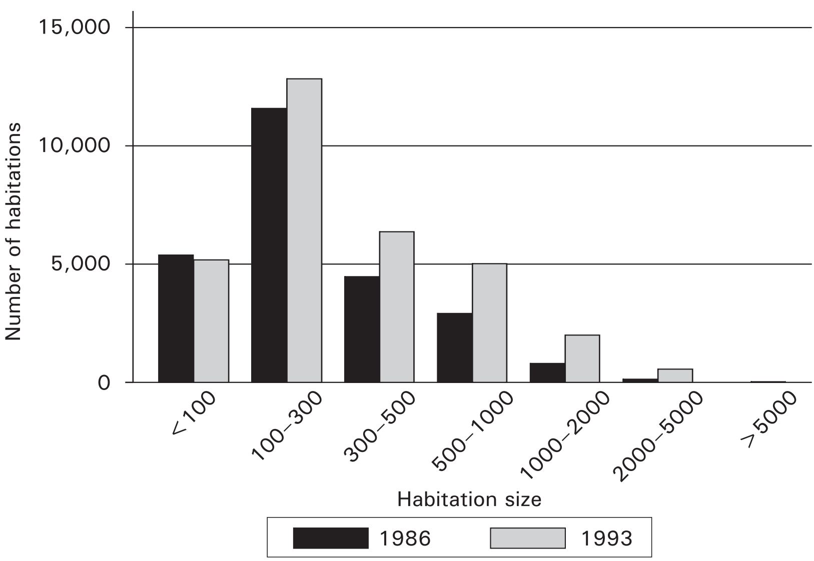
6. A school size of 100 or less is generally considered to be unviable.

FIGURE 1. Distribution of Habitations by Habitation Size, Uttar Pradesh, 1986–2002



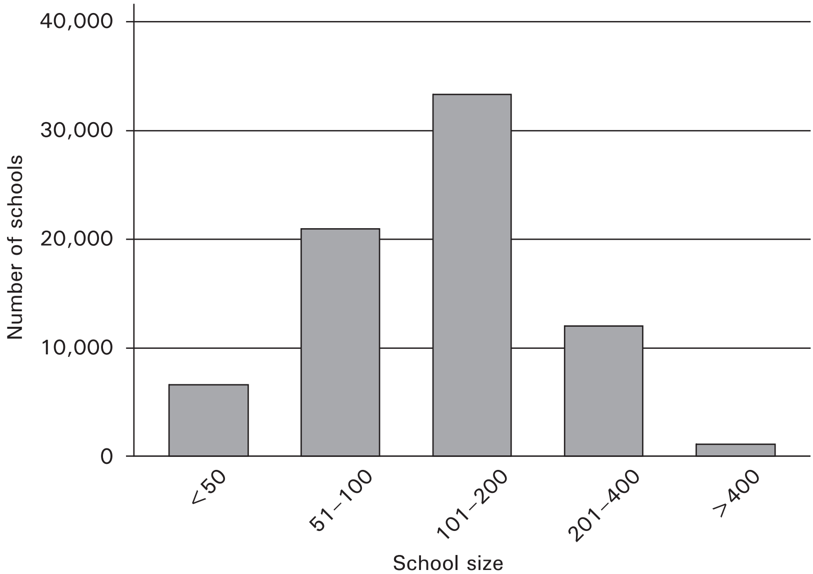
Source: AIES various rounds.

FIGURE 2. Distribution of SC and ST Habitations by Size, Uttar Pradesh, 1986 and 1993



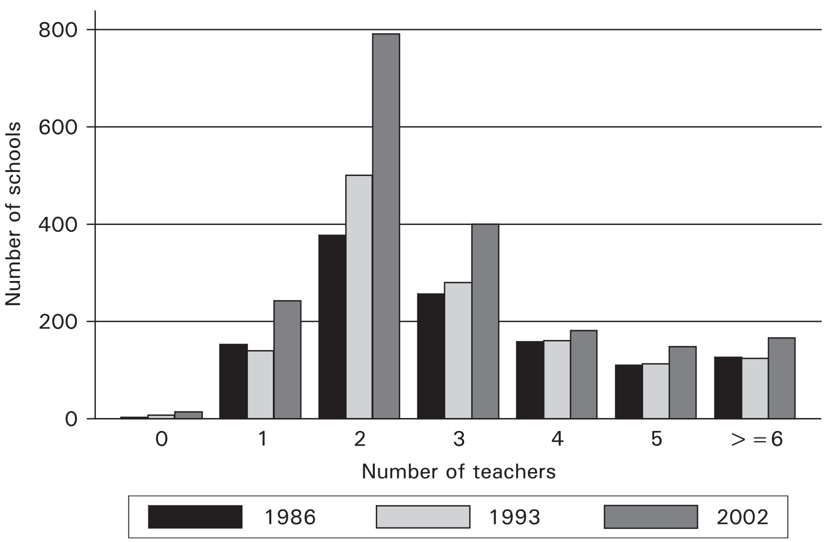
Source: AIES various rounds.

FIGURE 3. Distribution of Schools by School Size, Rural Uttar Pradesh, 1993



Source: AIES various rounds.

FIGURE 4. Distribution of Schools by Number of Teachers, Rural Uttar Pradesh, 1986–2002

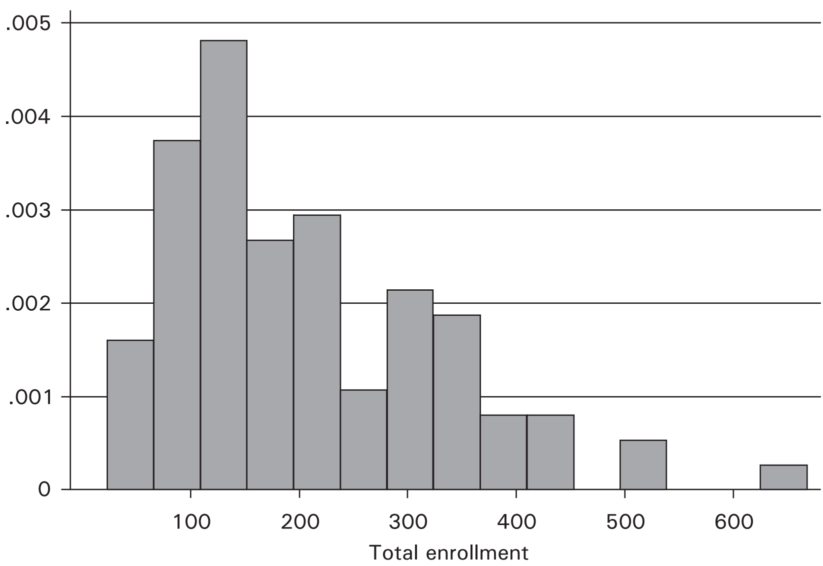


Source: AIES various rounds.

teachers, implying that multi-grade teaching, which combines students of several different grades, is the norm.[7]

The school-level data on enrollment and teachers in the seventh AIES reveal the considerable variation in school size (total enrollment) even within a block (figure 5). This variation confirms that the policy of using habitations as the basis for school mapping exercise in conjunction with the small size of habitations yields school populations which are not efficiently determined—if they were, there would be little variation in school size within this rather small geographical unit. In fact, large villages with several habitations in them are also characterized by significant variation in the size of schools located within a village. This is strikingly revealed in figure 6, where, for

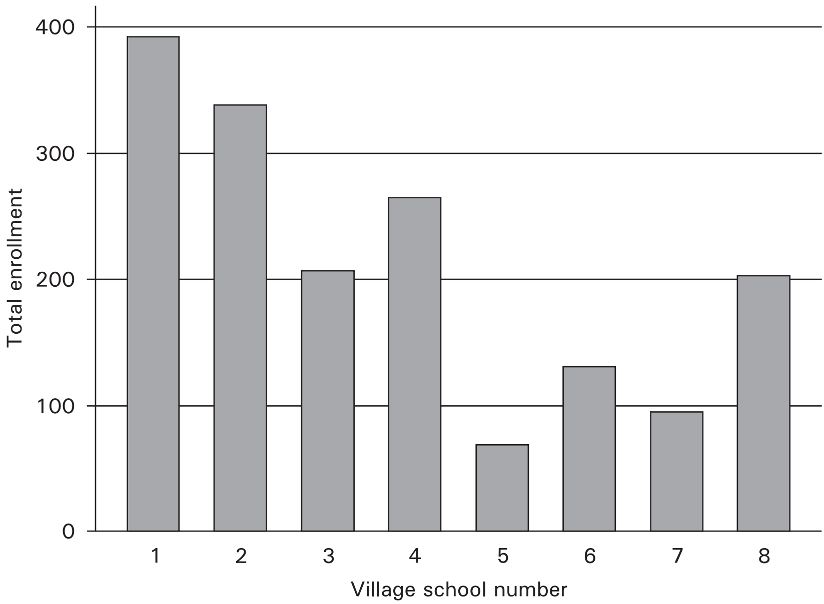
FIGURE 5. Frequency Distribution of Schools by Total Enrollment, Ballia District, Block Bairiya, Uttar Pradesh, 2002 (Mean Enrollment = 206)



Source: Seventh AIES round.

7. The effects of multi-grade teaching have not credibly been established, and many believe that its effects may be positive. However, there is agreement that the benefits are only obtained if it is restricted to two contiguous grades (for example, grades 1 and 2). In rural India, however, the assignment of teachers is based on class size, which is determined by the number of children of a particular age group in the habitation. It is thus common to see a large number of students in one class and a small number in other classes. Accordingly, the assignment of teachers to classes reflects class size, not an optimal allocation.

FIGURE 6. Distribution of 8 Schools by Size, Village ..., District Ballia, Uttar Pradesh



Source: Seventh AIES various rounds.

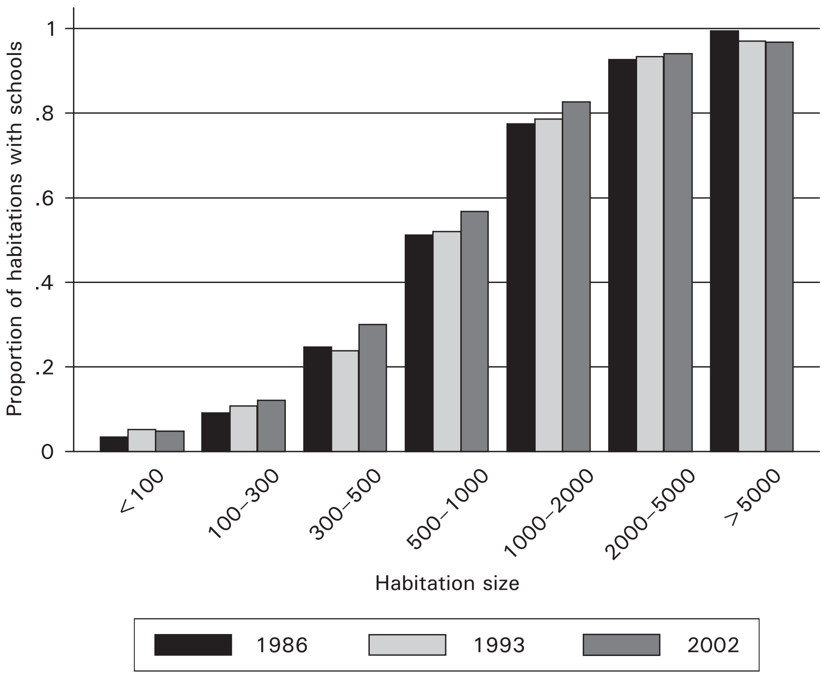
one village, I have plotted the distribution of school size across the eight schools located in the village. School size in this village ranges from less than 100 to nearly 400.

The data on the proportion of habitations with schools in them (figure 7) reveal that over 80 percent of habitations with a population of more than 1,000 have a primary school located in the habitation. Approximately half of the habitations with a population size of 500–1,000 have a school located in them. This number falls off sharply after that: in 1993, 24 percent of habitations with a population size of 300–500 had a primary school located in the habitation while the corresponding percentage for habitations of size less than 300 was only 6.9 percent.

The AIES does not provide school-level data on the caste composition of the student population.[8] Hence, it is not possible to show the extent of caste segregation in schools in Uttar Pradesh. However, the extensive residential segregation by caste, in separate habitations, suggests a correspondingly

8. For the seventh survey, data on enrollment is available at the level of individual schools, but no details on the caste composition of students are available.

FIGURE 7. Proportion of Habitations with Schools in Them, by Habitation Size, Uttar Pradesh, 1986–2002



Source: AIES various rounds.

high segregation of students by caste. This situation is not unique to Uttar Pradesh; since the school location policy is the one adopted by the Central Government, the same pattern of caste segregation across schools exists in other parts of India. The data from a survey of sixty schools in rural Andhra Pradesh (forty-five "main village" schools and fifteen schools located in the SC/ST hamlets) reveals the considerable caste segregation in this state.[9] In 2004, the proportion of SC/ST students to total students was as high as 66 percent in hamlet schools. Indeed, the SCs/STs and the OBC comprise almost all the student population (92 percent). In striking contrast, the SCs/STs comprise only 16 percent of the student population in "main village" schools. This survey also reveals the difference in size and correspondingly in teacher strength across hamlet and main village schools.

9. The first round of this survey was conducted in 2002, by the author, in collaboration with the Byrraju Foundation, Hyderabad.

The average (total) enrollment in hamlet schools in 2004 was ninety students while the average number of teachers was three. Conversely, main village schools had, on an average, an enrollment of 177 students, with an average of six teachers.

Data

For the analysis of this paper, I combine household data from the NSS on the schooling status of individuals, with district-level information on school availability and habitation characteristics from the AIES. Specifically, I examine the enrollment status of two cohorts of 6–11 year olds, from the 1993 and 1999 NSS employment surveys (rounds 50 and 55) and match this with data from the sixth (1993) and seventh (2002) AIES surveys on the proportion of habitations within a district with schools.

Each survey round of the AIES provides the basis for the construction of schools and the mapping programs in the inter-survey years. Thus, the construction of schools between 1986 and 1993, and hence the availability of schools in 1993, is based on the data contained in the fifth AIES survey (1986), while the availability of schools in 2002 is based on the sixth (1993) survey. As explained in detail in the next section, data from the preceding AIES survey thus provides the source of identification of the current availability of schools.

Matching of the NSS surveys and the AIES data is done at the level of the district. This exercise is complicated because, over the years, many districts have been divided into smaller units or new disricts have been created. Because we use data spanning the period 1986–2002, matching is done on the basis of the 1986 classification of districts. Information regarding the division of districts, which enables the tracking of districts over time, is provided in the Census of India.

Habitation Size and School Attributes

This section examines simple correlations between the characteristics of habitations and schools, focusing on the size of the habitation. The objective is to show a relationship between the attributes of habitation and quality of school q, as measured by the number of teachers and other features of schools. The next section provides evidence on the effect of these school features on enrollments, and allows an assessment of the extent to which the correlations between habitation size and schooling outcomes reflect habitation effects on school attributes.

In addition to school access, measured by the proportion of habitations in a district with a school located within the habitation, I examine the determinants of two other characteristics of the school system: the proportion of schools with two or less teachers and the probability of a school being located in a SC or ST habitation (given by the ratio of the number of SC/ST habitations with schools to total habitations). This latter variable indicates the degree of caste segregation in schools, since it is only when a school is available in a SC/ST habitation within the village that children from the SCs and the STs are likely to attend separate schools from children of other castes. The two characteristics of habitations which I initially focus on are the proportion of habitations with a population size of 500 or more and the average size of habitations in a district. The former variable is of particular interest, since it forms the basis of school location policies. The data are at the level of the district. The school attributes are from the sixth and the seventh AIES (1993 and 2002, respectively), and are related to habitation characteristics from the prior rounds of the AIES (fifth and sixth rounds for 1986 and 1993, respectively).

Table 1 presents simple correlation coefficients for these variables. The data reveal that districts characterized by relatively large habitations are more likely to have a school located in them. They are, however, also more likely

TABLE 1. Simple Correlation Coefficients between Habitation Size and School Attributes

	Sch_hab	*Prop schools with < =2 teachers*	*Sch_SChab*	*Hab_5*	*Avge pop.*
Sch_hab	1.0				
Prop schools w/ < =2 teachers	0.25	1.0			
Sch_SChab	0.57	0.17	1.0		
Hab_5	0.73	–0.17	0.31	1.0	
Avge pop.	0.70	–0.24	0.23	0.91	1.0

Note: Calculations are based on district-level data from the All India Education Surveys. Data on school attributes are from the sixth (1993) and seventh (2002) surveys, and are related to habitation size from the fifth (1986) and sixth (1993) surveys, respectively. Total number of observations is 110. Hab_5 is the proportion of habitations of population 500 or more. Sch_hab is the proportion of habitations with a school within the habitation, while Sch_SChab is the ratio of the number of SC/ST habitations with a school to total habitations. Avge pop. is the average population size of a habitation.

to have schools located in the SC/ST habitations, suggesting a greater degree of caste segregation in schools. Finally, larger habitations are associated with fewer schools with two or less teachers.

The table 2 elaborates on this evidence through OLS regressions which control for the effect of various other demographic and economic characteristics of habitations, including the average population and squared population of habitations in the district, average population (and its square) of SC/ST habitations, number of habitations and squared habitations separately for total habitations and SC/ST habitations, district female literacy rate, average per capita expenditure, and education rank. In addition to the proportion of

TABLE 2. Simple Regressions of School Attributes on Habitation Characteristics

	Pr sch_hab	*Pr. Tchrs < =2*	*Pr Sch_SChab*
Pr hab_5	0.65*	0.86	0.008
	(0.25)	(0.59)	(0.09)
Pr SChab_5	0.43	−1.39	0.69*
	(0.63)	(1.37)	(0.29)
Pr hab_3	−0.79	0.54	−0.29+
	(0.28)	(1.01)	(0.16)
Pr SChab_3	−0.46	−1.69	2.14*
	(0.88)	(3.03)	(0.77)
Avge pop.	−0.0007*	−0.001*	4.2 e−6
	(0.0003)	(0.0007)	(2.8 e−8)
Avge pop. square	1.69 e−7*	2.96 e−7	−1.49 e−8
	(7.47 e−8)	(1.86 e−7)	(2.77 e−8)
# of habitations	−0.0001*	−0.0002*	−0.00002*
	(0.00002)	(0.00007)	(7.2 e−6)
# of SC habitations	−0.00002	0.0003	−0.00002
	(0.00007)	(0.0003)	(0.00002)
District p.c. exp	−0.0001	−0.0002	−0.00003
	(0.001)	(0.0003)	(0.00005)
District ed rank	−0.0001	−0.0005+	0.00006
	(0.0001)	(0.0003)	(0.00005)
Lagged # of schools	0.0001*	0.001*	−1.32 e−6
	(0.0004)	(0.0001)	(0.00001)
Lagged enrollments	2.73 e−7	−4.6 e−6*	−5.79 e−9
	(2.74 e−7)	(9.1 e−7)	(1.08 e−7)
Dummy year 2	0.16*	−0.17+	0.04*
	(0.03)	(0.09)	(0.02)
Regression R^2	0.91	0.74	0.83
Sample size	110	110	110

Source: Computed by the author.

Notes: Standard errors are clustered at the district level.

*Significant at 5% level.

+Significant at 10% level.

total and SC/ST habitations of size greater than 500, regressions include the proportion of total habitation and SC/ST habitations with a population of 300–499, since school location policies also call for schools to be located in habitations of this size, if no school is available within a 1 km distance. As before, all regressions are run on district-level data. Because outcomes for any given district are likely to be correlated, all standard errors are clustered at the level of the district.

Supporting the evidence from simple correlation coefficients, the proportion of habitations of size 500, and more is positively and significantly associated with the proportion of habitations with schools, even after controlling for the average size of habitations (and its square). This suggests that policy rules *do* affect school location, since the rule dictates that particular population cut-off levels must determine school location. Interestingly, this same variable does not affect the availability of schools in SC/ST habitations—access to schools in these habitations reflects the size distribution of SC/ST habitations, as measured by the proportion of SC/ST habitations of population between 300 and 499, and greater than 500. The population cut-off values that determine school location also do not appear to affect teacher allocations: proportion of habitations of size between 300 and 499, and greater than 500 is not a significant determinant of the proportion of schools with two or less teachers. This provides supportive evidence that these population cut-off values *do* reflect policy rules; if they merely reflected a non-linear effect of habitation population size, they would also likely affect other schooling inputs. Teacher strength *is* significantly affected by the average population size of habitations: as expected, the proportion of schools with two or less teachers falls as population size increases.

Empirical Methodology for Assessing the Effect of School Characteristics on Enrollments

Basic Equation for Examining the Effect of Access to School

The basic estimating equation for assessing the effect of access to schools on the enrollment status of child *i* in district *j* is the following:

$$(1) \qquad \Pr(enroll)_{ij} = \alpha_0 + \alpha_1 (Sch\_hab)_j + X'_{ij}\alpha_2 + Z'_j\alpha_3 + u_{ij}$$

The dependent variable in this regression, $\Pr(enroll)_{ij}$, is an indicator variable which takes the value 1 if the child is currently enrolled in primary school, 0 otherwise. The regressor of interest, Sch_hab_j is the proportion of habitations in the district with a school located within the habitation. X_{ij} is a vector of child and household specific characteristics that determine school enrollment. This includes the age and squared age of the child, his or her gender, a dummy variable which records whether he or she is a SC/ST, household per capita expenditure, the demographic composition of the household as reflected in the division of household members across ten sex-age groups, the maximum years of education of adult household males in two different age groups (20–40 and 40–60), and an indicator variable for whether the household head has completed primary education. Z is a vector of district variables, including the district average per capita expenditure, female literacy rates, the number of habitations, and the proportion of the SC and the ST to total population in the district.

A primary concern of this paper is to examine the differential effects of school location policies on the SC and the ST. Therefore, all regressions are run on a pooled sample of children in the 6–11 age group and then separately for children who belong to SCs and STs, and those who belong to general castes.

Identification of School Access

The critical issue in the estimation of equation (1) relates to the identification of the effect of school availability on enrollment. Decisions relating to the location of schools are based on the size of the habitation. But habitation size will also determine total enrollments in the school and hence the number of teachers and a number of other attributes of a school. Large habitations may also have better infrastructure such as roads which are also likely to increase the returns to schooling. If so, α_1 in equation (1) will be biased upwards.

I base identification on the program rules which specify a specific population cut-off level for decisions regarding school location. Specifically, only habitations with a population exceeding 300 are eligible for schools. However, the criterion for determining the eligibility of habitations of sizes 300–499 differs from those of larger habitations, in that it includes a distance criterion (no other school within a 1 km distance) as well as a population criterion. Thus, population cut-off levels of 300 and 500 will differentially affect access to school.

These population cut-off levels are unique to decisions regarding location of school—they do not influence decisions regarding the availability of any other local public good, including health centers, roads, or investments in sanitation. This suggests the feasibility of basing identification on this particular size classification of habitations in a district, specifically on the proportion of habitations of size greater than 300 and greater than 500. Because the AIES is conducted specifically for the purpose of school mapping, it provides information on the number of habitations in a district which fall into these population groupings.

As previously described, data from a specific survey year formed the basis for the school mapping exercises conducted in the inter-survey period between that and the next survey. Thus, identification of the availability of schools in 1993, for example, reflects data on the distribution of habitations in 1986. It is these lagged population figures, the exact data which guided school mapping exercises, which are used in this analysis.

Habitation size, however, also determines school size. Moreover, data on the availability of schools in habitations, previously summarized in figure 7, suggest that targets for construction of schools were not achieved—not all habitations of population size greater than 500 have a school located in them. This reflects the financial constraints faced both by the Central Government and the state government—the necessary funds required to provide the desired number of schools have not been available. Instead, governments have had to prioritize amongst habitations in determining school location, making decisions on the basis of the district's educational ranking, as discussed in the second section of this paper.

The availability of a school in a habitation thus reflects the combined effect of the population criterion and the district's educational rank. This suggests the use of the interaction of the proportion of habitations of size 300–499 (*hab_3*) and greater than 500 (*hab_5*) with the district's education rank (*ed_rank*) as instruments. This educational index provides a particularly good source of identification, since it ranks all districts in the country based on 1991 data, not just districts within a given state, and because the particular rank of any given district was meaningful—the division of central funds for schooling was to be in proportion to the specific cardinal rank of the district. Moreover, the district's education rank amongst all districts in the country was not used to guide any state or local government decisions regarding schooling investments.[10]

To ensure that identification comes only from the interaction of the population criterion with the district's education rank, *hab_3* and *hab*_5 as well as

10. This ranking is available in Annexure V of the Report of the Working Group on Early Childhood Education and Elementary Education set up for Formulation of the Eigth Five Year Plan.

ed_rank are included amongst the set of regressors and hence excluded from the instrument set. I also include the average population size of habitations in the district amongst the regressors, as further assurance that identification does not come from habitation size.

Assessing the Implications of School Size

As previously noted, the decision to locate a school in a habitation suggests that the socio-economic features of the habitation, including habitation size, will affect schooling outcomes through their effects on school attributes such as school size and the composition of the student population. This section discusses the methodology used to evaluate how habitation size affects enrollments in schools.

Rather than consider the effect of school size, I assess the impact of the number of teachers in the school, an attribute which closely reflects school size. School size will primarily be determined by the population size of the habitation. However, habitation population may also affect other school inputs, including the number of teachers, as also the availability of other local public goods, such as health centers, which may have spillover effects on schooling. Lack of good instruments for school size therefore dictates my decision to consider, instead, the number of teachers.

The number of teachers assigned to a school reflects several criteria discussed in the second section of this paper. One criterion is total school enrollment. However, this rule does not generate a good source of identification. It implies that availability of teachers is determined by the size of the habitation which, as previously argued, is likely to be directly correlated with schooling outcomes for a variety of reasons. Instead, I base identification on the auxiliary set of rules specified under Operation Blackboard and the extension of this program in 1992, which determine the minimum number of teachers in a school. Under the expanded Operation Blackboard, each school with an enrollment of over 100 was to have three teachers (with funding provided by the Central Government). This implies a distinct population cut-off that determines whether the school would have two or fewer teachers: Two (or fewer) teacher schools would be those with a total enrollment of less than 100.

Identifying the number of teachers on the basis of this strategy requires information on the proportion of schools in a district with an enrollment of less than 100. Unfortunately, that information is not available in the AIES. However, the information necessary to predict this number is. Specifically, we assume that the distribution of school size follows a normal distribution, so that the proportion of schools of size less than 100 is:

$$(2) \qquad \Pr(Size <= 100) = \int_0^{100} \frac{1}{\sqrt{2\pi\sigma}} e^{-(100-\mu)^2/2\sigma^2}$$

In all probability the variation in school size will vary with the distribution of habitation size. I therefore assume that $\sigma = f(hab\_3)$. A linear approximation of this, inserted into equation (2), implies that the probability of school size being less than 100 is a function of $Size\_100 = -[(100 - \mu)/(hab\_3)]$, where μ is the (lagged) average enrollment in the school.

I use $(Size\_100)^2$ and $(Size\_100)^3$ as instruments for the proportion of schools with two or less teachers. Identification thus comes from a non-linear function of the cut-off enrollment size which determines this distribution of teachers (100), lagged average enrollment in the district, and the proportion of habitations of size 300–500 in the district, with the particular functional form suggested by the probability distribution of school size. Lagged average enrollments and habitation size are from data in the survey round of the AIES previous to the NSS survey round from which the current enrollment status of the student is drawn. That is, for children in the 6–11 age group in 1993, lagged average enrollments and habitation size are from the fifth (1986) round of the AIES.

As before, stronger identification comes from interacting $(Size\_100)^2$ and $(Size\_100)^3$ with the district's educational index rank. Because teachers provided under Operation Blackboard and the expanded version of this scheme were funded by the Central Government, districts which had the highest rank were most likely to receive these funds on a priority basis. The interacted variable, $(size\_100)^2*ed\_rank$ and $(size\_100)^3*ed\ rank$ ensures that identification does not just come from a non-linear function of (lagged) average enrollments and habitation size.

Assessing the Implications of Residential Segregation

A final concern is that the government's school location policies translate the residential segregation which characterizes rural India into a *de facto* system of caste-based schooling segregation, and that this provides an additional avenue by which school policies affect schooling attainment. My analysis of this issue draws on the insight that schooling segregation follows when schools are provided in SC/ST sub-habitations of a village. When this is not the case, then the SC/ST students must attend the same schools as students from the other castes. This suggests that the extent of schooling segregation in a district will reflect the availability of schools in the SC/ST habitations.

I therefore examine the effect of the probability of a school being located in a SC/ST habitation on enrollment decisions. This probability is given by the joint probability of a school being located in a habitation and the habitation being a SC/ST habitation, whose sample counterpart is the ratio of the number of SC/ST habitations with a school to the total number of habitations in the district. I refer to this variable as (*Sch_SChab*).

The analysis of the effect of the availability of schools in the SC/ST habitations, as distinct from that of overall access to schools, is possible because the AIES provides data on the total number of SC/ST habitations, their population distribution, and the number of them that have a school located in them. As in the treatment of the endogeneity of the (general) availability of schools (*Sch_hab*), my instruments for (*Sch_SChab*) are formed by interacting the district's educational rank with the number of SC/ST habitations of size 300–499 and the proportion of SC habitations with a population of 500 or more. These two population characteristics of SC/ST habitations, *SChab_3 and SChab_5*, are included amongst the regressor set. I also expand the regressor set to include the proportion of the SC/ST habitations in the district and the average size of SC/ST habitations.

The expanded regression that I run is:

$$\Pr(enroll)_{ij} = \alpha_0 + \alpha_1(Sch\_hab)_j + X'_{ij}\alpha_2 + Z'_j\alpha_3 + \alpha_4(Tch\_2)_j + \alpha_5(Sch\_SChab)_j + u_{ij} \quad (3)$$

where $(Tch\_2)_j$ is the district proportion of schools with two or less teachers.

The first-stage reduced form regressions for (*Tch_2*) and (*Sch_SChab*) include the vector of variables, which capture the population distribution of habitations in the district (*hab_3, hab_5, SChab_3, SChab_5*) as well as the interaction of these variables with the district's educational rank. Thus, the first stage regressions provide evidence of the effect of the size of the habitation on schooling outcomes. Combining this with the second-stage estimates of the effect of school attributes on enrollment outcomes [the coefficients in equation (3)], it is possible to examine the extent to which habitation characteristics affect schooling outcomes through their effect on school characteristics, independent of any direct effect through other means.

Results

Preliminary Regressions on School Access

Table 3 reports results from an initial set of regressions, which considers the effect of the availability of a school on enrollment decisions in a habitation (equation 1). The first column reports the reduced form first-stage

TABLE 3. Effect of the Availability of a School, in Habitation, on Enrollment in Primary School, Rural Uttar Pradesh, Children Ages 6–12

	Proportion of habitations with schools	*Currently enrolled in primary school*		
		Probit	*IV probit*	*Marginal effect*
Prop. Habitations with schools	–	0.83* (0.13)	2.28* (0.35)	0.48
Habitations 300–500 × rank	–8.87 e–7* (2.14 e–7)	–	–	–
Prop. Habitations greater than 500 × rank	–0.002* (0.0005)	–	–	–
Habitations 300–500	0.0001 (0.0001)	0.0001* (0.00006)	0.0003* (0.0001)	0.00006
Prop. Habitations greater than 500	0.89* (0.17)	–0.70* (0.18)	1.495* (0.256)	0.32
SC/ST	0.005 (0.003)	–0.03 (0.02)	–0.031 (0.025)	–0.007
Sex	–0.001 (0.001)	–0.34* (0.03)	–0.34* (0.03)	–0.072
Age	–0.004 (0.004)	1.43* (0.08)	1.43* (0.08)	0.302
Household p.c. exp	–2.57 e–6 (4.45 e–6)	0.0002* (0.00006)	0.0002* (0.0006)	0.00004
District avge p.c. exp	–0.0001 (0.0002)	–0.00007 (0.0002)	–0.0001 (0.0002)	–0.00002
Educational rank	0.001* (0.0004)	–0.0004* (0.0002)	–0.0005* (0.0002)	–0.0001
# of habitations	–0.00005* (0.00002)	0.00002 (0.00002)	0.00004+ (0.00002)	0.00001
Avge pop. of habitations	0.00002 (0.00004)	–0.00006 (0.00006)	–0.0001+ (0.00006)	–0.00002
District Prop. SC	0.026 (0.12)	–0.04 (0.19)	–0.16 (0.20)	–0.034
Sample size	16,759	16,759	16,759	
Test Statistic	$F(29,61) = 104.14$	$\chi^2 = 2004.48$	$\chi^2 = 1843.64$	
Wald test of exogeneity $\chi^2(1)$			20.35 ($p = 0.00$)	

Notes: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

+Significant at 10% level.

regression of the proportion of habitations with schools in them on the set of instruments. All regression errors are clustered at the district level, to account for the correlation between observations within a given district. The second column reports results from a probit regression of school enrollment on (*Sch_hab*), ignoring the endogeneity of this variable. The third column uses an instrumental-variable probit (IV probit) regression, with suitably adjusted standard errors, to account for endogeneity, using the instruments previously described. The last column lists the marginal effects of the variable on the probability of current enrollment in primary school, based on the IV probit results of the previous column.

The first stage regression attests to the explanatory power of the instruments. Larger habitations (with population in excess of 500) are more likely to have a school in them. However, the effect of population size on school access is tempered by the district's educational rank in the expected direction. The government policy specified that districts with a *lower* educational rank should first be allotted schools. Reflecting this, the interaction of habitation size with educational rank negatively affects the proportion of habitations in the district with a school. That is, a district with sufficiently large habitations was less likely to get a school if its educational rank was high.

The coefficient on school access obtained from probit regressions which ignore the endogeneity of school location policies is significantly lower than that obtained from IV probit regressions, which utlize the interaction of the variables used to guide policy with the education rank of the district as instruments. Correspondingly, Wald tests for the exogeneity of school access reject the hypothesis of no endogeneity. These results suggest that prior estimates of a small effect of distance to school on enrollment decisions may be biased downwards because of the failure to properly account for the endogeneity of school access. The IV probit results suggest a relatively large effect of the availability of schools in a habitation. The coefficient of 2.28 from the IV regression results reported in table 3 imply that a 1 percent increase in the proportion of habitations with a school will increase the proportion of school-age children attending schools by 0.4 percent. This justifies the decision to situate a school within easy access to students, if this policy is to be evaluated purely from the viewpoint of its effect on school access.

Table 4 provides estimates of school access by caste and by gender. While school access is important for children from all castes, and for boys and girls, it matters more for SCs and for girls. The effect on enrollment of the probability of a school being located in a habitation is 3.39 for the SC/ST but only 1.94 for the upper castes. This translates into an elasticity of enrollment

TABLE 4. Effect of School Access by Caste and Gender

(Dependent variable: currently enrolled in primary school, ages 6–12)

	SC/ST	*Other castes*	*Boys*	*Girls*
Prop. Habitations with schools	3.39*	1.94*	1.81*	2.86*
	(0.81)	(0.38)	(0.46)	(0.53)
Habitations 300–500	0.0002+	0.0003*	0.0002+	0.0005*
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Prop. Habitations > 500	–2.41*	–1.31*	–1.20*	–1.92*
	(0.57)	(0.29)	(0.35)	(0.38)
SC/ST	–	–	–0.05+	–0.01
			(0.03)	(0.04)
Sex	–0.43*	–0.32*	–	–
	(0.05)	(0.03)		
Age	1.08*	1.54*	1.49*	1.39*
	(0.16)	(0.09)	(0.11)	(0.12)
Hhold p.c. exp.	0.0004*	0.0002*	0.0001	0.0004*
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
District p.c. exp.	–0.0003	–0.0001	0.0002	–0.0004
	(0.0004)	(0.0002)	(0.0003)	(0.0003)
Education rank	–0.0002	–0.0005*	–0.0005*	–0.0004*
	(0.0004)	(0.0002)	(0.0003)	(0.0002)
# of habitations	0.0001	0.00005+	0.0001*	0.00002
	(0.0001)	(0.00003)	(0.00003)	(0.00004)
Avge habitation population	–0.0001	–0.0001+	–0.0002+	–0.00007
	(0.0001)	(0.00007)	(0.00009)	(0.0001)
Prop. SC	–0.05	–0.22	–0.18	–0.10
	(0.37)	(0.23)	(0.27)	(0.29)
Sample size	3,863	12,896	9.143	7,616
Wald χ^2 (27)	425.82	1456.43	804.64	997.89
Wald test of exogeneity, $\chi^2(1)$	8.28	12.06	8.25	12.87
	(p=0.004)	(p=0.001)	(p=0.004)	(p=0.0003)

Notes: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

+Significant at 10% level.

with respect to school access of 0.7 for the SC/ST children, but only 0.3 for the other castes. Similarly, the corresponding coefficient of 1.81 for boys and 2.86 for girls implies an elasticity of 0.3 for boys but 0.5 for girls.

Incorporating Teacher Availability and School Segregation

I now extend the analysis to see how enrollment decisions are affected by additional school characteristics, and to examine whether these effects vary by caste. Table 5 documents results from the first stage regressions

TABLE 5. First Stage Regressions

	Proportion habitations with school	*Proportion of schools with < =2 teachers*	*SC habitations with schools to total habitations*
Habitations 300–499 × rank	–1.51 e–6*	–4.73 e–6*	–2.94 e–7*
	(4.43 e–7)	(1.10 e–6)	(1.26 e–7)
Prop. Habitations > 500 × rank	–0.002*	–0.001	0.0003+
	(0.0008)	(0.002)	(0.0002)
Predicted prop. of schools with < =2 teachers	–1.37 e–7	4.39 e–6*	–1.35 e–7
	(4.27 e–7)	(1.17 e–6)	(8.18 e–8)
Predicted < =2 teachers × rank	3.81 e–9	–1.56 e–8*	1.07 e–9
	(3.05 e–9)	(7.88 e–9)	(7.37 e–10)
Predicted prop. < =2 teachers, square	–2.15 e–11	2.84 e–9*	–2.68 e–11
	(5.21 e–10)	(1.34 e–9)	(8.83 e–11)
Predicted prop. < =2 teachers square × rank	2.09 e–12	–1.20 e–11	1.43 e–13
	(3.54 e–12)	(9.00 e–12)	(6.75 e–13)
SC habitations 300–500 × rank	7.36 e–6	0.00001+	1.43 e–6
	(2.85 e–6)	(8.29 e–6)	(8.75 e–7)
SC habitations > 500 × rank	–0.006	–0.04*	–0.003*
	(0.006)	(0.02)	(0.001)
Habitations 300–499	0.0001+	0.001*	0.00002
	(0.00006)	(0.0002)	(0.00002)
Prop. habitations > 500	0.292	–1.04+	–0.04
	(0.252)	(0.62)	(0.05)
SC habitations 300–499	–0.001*	–0.001	–0.0001*
	(0.0002)	(0.001)	(0.00007)
SC habitations > 500	3.71*	5.16	1.50*
	(1.06)	(4.27)	(0.37)
# of habitations	–0.0001*	–0.0001	–2.52 e–6
	(0.00002)	(0.0001)	(4.68 e–7)
Avge hab. population	0.0002	0.001*	3.71 e–6
	(0.0001)	(0.0002)	(0.00003)
District prop. SC	–1.10*	2.90	0.06
	(0.53)	(2.47)	(0.19)
Sample size	16,759	16,759	16,759
Regression R^2	0.92	0.74	0.90

Note: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

+Significant at 10% level.

of the endogenous variables on the full set of instruments and additional regressors.

The regression results confirm the explanatory power of the instruments, and hence the role of habitation size in determining important school attributes. For example, the interaction of a district's education rank with the proportion of habitations of size greater than 500, significantly affects the proportion of habitations with a school. And, the interaction of $(size\_100)^2$ and $(size\_100)^3$ with education rank also significantly reduces the proportion of schools with less than two teachers: As prescribed by policy, a greater proportion of schools with enrollment over 100 suggested additional teachers, and hence a reduction in the proportion of schools with two or fewer teachers at the time of the next survey. However, this effect is stronger in districts with lower educational rank.

Confirming the results from simple correlation coefficients, the first stage regressions also suggest that the instruments for the proportion of schools with two or less teachers do not also affect access to schools. This supports the credibility of the results, since it suggests that the instruments for teacher strength are uncorrelated with the availability of schools and other schooling inputs.

These first stage regression results provide the basis for the IV probit regressions in table 6. The first column repeats regression results for the probability of current enrollment in primary schools as a function of school access, ignoring other school attributes, to provide a point of comparison.[11] The regression reported in column (2) includes the proportion of schools with two or less teachers and the probability of a school in a SC/ST habitation amongst the regressors. The results confirm that school access increases enrollment. However, as expected, fewer teachers imply a reduction in enrollment. And, higher levels of school segregation, suggested by the availability of schools in SC/ST habitations, appear to increase enrollments.

These results in conjunction with the first stage regression results of table 5 suggest that habitation size does affect schooling outcomes because of school location policies which determine whether a school can be constructed in a particular habitation, including the availability of schools in the SC/ST habitations. Habitation size also affects school outcomes through its effect on other schooling inputs such as the number of teachers. This is because the assignment of teachers to schools is based on enrollment, which, in turn, reflects the population of the habitation in which the school is located. That the effects of habitation size on schooling outcomes are mediated through school inputs is further suggested by comparing the results in column 1, which conditions only on school access, with those in column 2 which additionally control for teacher strength and the provision of schools in the SC/ST communities. The inclusion of these latter variables significantly affects the coefficient on measures of habitation size which are included as regressors in both regressions.

11. This regression differs from that reported in table 3, since it includes additional regressors.

TABLE 6. Effect of School Access Controlling for Number of Teachers and School Segregation

	Probability of current enrollment in primary school		
	(1)	*(2)*	*(3)*
Prop. Habitations with schools	2.34*	3.59*	3.42*
	(0.39)	(0.42)	(0.44)
Prop. of schools with < =2 teachers	–	–0.35*	–0.23^{+}
		(0.07)	(0.13)
Prob. of school in SC habitation	–	4.59*	4.93*
		(1.58)	(1.61)
Habitations 300–500	–0.001*	0.0001	0.0001
	(0.0003)	(0.0001)	(0.0001)
Prop. Habitations greater than 500	–3.05*	–1.31*	–0.97*
	(1.61)	(0.26)	(0.39)
SC habitations 300–500	0.001*	0.001*	0.001*
	(0.0003)	(0.0003)	(0.0003)
Prop. SC habitations >500	–3.05^{+}	–11.02*	–10.43*
	(1.61)	(2.23)	2.38
SC/ST	–0.03	–0.06*	–0.06*
	(0.02)	(0.03)	(0.03)
Sex	–0.34*	–0.34*	–0.34*
	(0.03)	(0.03)	(0.03)
Age	1.4	1.45*	1.45*
	4*	(0.08)	(0.08)
	(0.08)		
Household p.c. exp	0.0002*	0.0002*	0.0002*
	(0.00006)	(0.0001)	(0.0001)
District avge p.c. exp	0.0002	0.0004*	0.0005*
	(0.0002)	(0.0002)	(0.0002)
Educational rank	–0.0003	–0.0009*	–0.0008*
	(0.0002)	(0.0002)	(0.0002)
# of habitations	0.0001*	0.0002*	0.0001*
	(0.00003)	(0.0003)	(0.00003)
# of SC habitations	–0.0002*	–0.001*	–0.0005*
	(0.0001)	(0.0001)	(0.0001)
Avge pop. of habitations	–0.0002^{+}	–0.0003*	–0.0005*
	(0.0001)	(0.0001)	(0.0002)
Avge pop. SC habitations	–0.0002	–0.0003^{+}	–0.0003^{+}
	(0.0001)	(0.0002)	(0.0002)
District prop. SC	–1.08	0.35	0.03
	(0.95)	(1.02)	(1.06)
Predicted prop. teachers< =2	–	–	–3.31 e–7
			(4.07 e–7)
Predicted prop. teachers < =2 squared	–	–	–9.34 e–11
			(2.62 e–10)
Sample size	16,759	16,759	16,759
Wald χ^2	1871.72	1883.91	1884.43
Wald test of exogeneity χ^2	17.57	91.02	87.36

Notes: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

$^{+}$Significant at 10% level.

Robustness Checks

Before assessing the separate implications of schooling inputs by caste, I conduct several robustness checks. First, I include $(size\_100)^2$ and $(size\_100)^3$ amongst the regressors, allowing identification of teacher strength to come only from the interaction between these variables and the district's education rank. This is a simple over-identification test for the validity of including these variables in the instrument set. The results, detailed in the last column of table 6, support their validity: their independent effect on school enrollment is statistically insignificant and does not change the regression results.

A second check explores the sensitivity of results to non-linearity in habitation size. As always, when identification rests on a particular functional form assumption, there is the possibility that the results simply reflect a non-linear relationship between habitation size and school enrollment. To test for this, a second set of regressions includes the square of average habitation population as well as the square of the average population of SC/ST habitations in the regressor set. The results, reported in table 7, suggest that the results are robust to this inclusion. The estimates of the effect of school inputs, including school access, are not affected by this inclusion. And, standard F tests for the expanded regression reject this version in favor of the more parsimonious specification.

Regression Results by Caste

In order to examine whether the effects of school attributes varies by caste, I run the same regression on schooling enrollments separately for children from the SC and the ST and for those from the other castes. The results are reported in table 8. For all children, regardless of caste, access to schools enhances the probability of enrollment while the proportion of schools with two or fewer teacher reduces this probability. However, there are significant differences in the effect of schooling segregation. An increase in the probability of a school being located in a SC/ST habitation significantly enhances the schooling of the upper caste households. Conversely, the effect on children from the SC and ST is negative. However, this latter effect, though relatively large in magnitude, is imprecisely estimated and is not significant at conventional levels of significance.

The results suggest that a greater degree of homogeneity in the student population enhances the schooling attainment of upper caste households. It is beyond the scope of this paper to explore the reasons that underlie this effect. One explanation offered in the literature is that ethnic homogeneity fosters community support for, and involvement in, schools, thereby improving schooling attainment (Miguel and Gugerty, 2005). However, if this were

TABLE 7. Robustness Check: Including Square of Average Population of Habitation

	First stage regressions			
	Prob. school in habitation	*Prop. teachers < =2*	*Prob SC hab with school*	*Current enrollment*
Prob. School in habitation	–	–	–	3.87* (0.44)
Prop. Teachers < =2	–	–	–	–0.32* (0.06)
Prob. School in SC habitation	–	–	–	2.77* (1.44)
Habitations 300–499 × rank	–1.39 e–6* (4.18 e–7))	–3.74 e–6* (9.60 e–7)	–3.31 e–7* (1.24 e–7)	–
Prop. Habitations > 500 × rank	–0.002** (0.0007)	–0.0001 (0.002)	0.0003+ (0.00018)	–
Predicted prop of schools with < =2 teachers	–2.64 e–7 (4.65 e–7)	3.3 e–6* (9.56 e–7)	–8.37 e–8 (8.99 e–8)	–
Predicted < =2 teachers × rank	4.54 e–9 (3.28 e–9)	–9.33 e–9 (6.57 e–9)	7.89 e–10 (7.41 e–10)	–
Predicted prop < =2 teachers, square	–3.3 e–10 (6.64 e–10)	1.82 e–10 (1.08 e–9)	9.27 e–11 (1.12 e–10)	–
Predicted prop < =2 teachers square × rank	3.60 e–12 (4.23 e–12)	1.02 e–12 (7.14 e–12)	–4.24 e–13 (7.54 e–13)	–
SC/ST habitations 300–500 × rank	7.19 e–6* (2.92 e–6)	0.00001* (6.02 e–6)	1.46 e–6+ (8.18 e–7)	–
SC/ST habitations > 500 × rank	–0.005 (0.005)	–0.03 (0.02)	–0.003* (0.001)	–
Habitations 300–499	0.0001 (0.0001)	0.0004 (0.0002)	0.00002 (0.00002)	0.0002 (0.0001)
Prop. habitations > 500	0.48* (0.23)	0.54 (0.42)	–0.11* (0.05)	–1.17* (0.41)
SC/ST habitations 300–499	–0.001* (0.0003)	–0.0007 (0.0006)	–0.0001+ (0.00007)	0.001* (0.0003)
SC/ST habitations > 500	3.70* (1.06)	5.00 (3.96)	1.49* (0.34)	–8.00* (2.12)
# of habitations	–0.00005* (0.00002)	–0.0001 (0.0001)	–3.27 e–6 (4.6 e–6)	0.0002* (0.00003)
# of SC/ST habitations	0.0002* (0.00005)	0.0002 (0.0001)	–7.08 e–6 (0.00001)	–0.0005* (0.0001)
Avge hab. population	–0.0002 (0.0004)	–0.003* (0.0007)	0.0002* (0.0001)	–0.001* (0.0005)
Avge hab. population square	1.40 e–7 (1.24 e–7)	1.21 e–6* (2.42 e–7)	–6.15 e–8* (2.95 e–8)	3.75 e–7* (1.42 e–7)
Avge pop. SC/ST habitation	–0.00001 (0.0003)	–0.0001 (0.001)	–0.0001 (0.0001)	0.001* (0.0005)
Avge SC/ST pop. square	–6.62 e–10 (1.68 e–7)	–2.64 e–8 (4.11 e–7)	2.39 e–8 (3.07 e–8)	7.84 e–7* (2.35 e–7)
Sample size	16,759	16,759	16,759	16,759
Regression R^2	0.92	0.80	0.90	Wald $\chi^2 = 1893.05$

Notes: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

**Significant at 1% level.

+Significant at 10% level.

TABLE 8. Enrollment Regressions by Caste

	SC/ST		Other castes	
Prop. Habitations with schools	4.30*	(0.87)	3.24*	(0.48)
Prop. of schools with < = 2 teachers	–0.21+	(0.13)	–0.35*	(0.08)
Prob. of school in SC habitation	–2.34	(3.00)	6.45*	(1.90)
Habitations 300–500	–0.0002	(0.0002)	0.0002	(0.0001)
Prop. Habitations greater than 500	–1.59*	(0.51)	–1.25*	(0.31)
SC habitations 300–500	0.0005	(0.0006)	0.001*	(0.0003)
Prop. SC habitations > 500	–4.48	(4.77)	–12.67*	(2.70)
Sex	–0.43*	(0.05)	–0.32*	(0.03)
Age	1.09*	(0.07)	1.55*	(0.09)
Household p.c. exp	0.0005*	(0.0002)	0.0002*	(0.0001)
District avge p.c. exp	0.0005	(0.0004)	0.0003+	(0.0002)
Educational rank	–0.0002	(0.0005)	–0.0009*	(0.0003)
# of habitations	0.0002*	(0.0001)	0.0001*	(0.00004)
# of SC habitations	–0.0004+	(0.0002)	–0.0005*	(0.0001)
Avge pop. of habitations	–0.00004	(0.0002)	–0.0004*	(0.0001)
Avge pop. SC habitations	–0.001*	(0.0003)	–0.0001	(0.0002)
District Prop. SC	2.91	(2.07)	–1.23*	(0.31)
Sample size	3863		12896	
Wald χ^2	448.71		1480.90	
Wald test of exogeneity χ^2	18.21	(0.0004)	70.87	(0.0000)

Note: Regressions run on data from the 55th (1999) and 50th (1993) rounds of the National Sample Surveys. All standard errors are clustered at the district level.

*Significant at 5% level.

+Significant at 10% level.

so, the greater ethnic homogeneity fostered by schooling segregation should increase schooling enrollments for the SCs/STs as well as for other castes. A more convincing explanation is that schooling attainment does depend on the "mean" quality of the community, either through its influence on the attributes of the student population or through the quality of community contributions to the schools. If so, and if the SC/ST communities are characterized by lower mean quality then, consistent with the results of this paper, schooling segregation will enhance the schooling of the upper castes while reducing it for the SCs and the STs.

Quantifying the Effects of Habitation Size on Enrollment through School Characteristics

Combining results from first stage regressions, which reveal the influence of school location policies on schooling inputs, with the results from the second stage instrumental variable regressions, which confirm that these same schooling inputs influence enrollment, establish that school location policies affect schooling outcomes not just because they affect a household's access to school but also because they affect schooling inputs. Since access to schools is greater in larger habitations, one would expect children who reside in larger habitations to have an advantage over their counterparts living in smaller habitations. However, habitation size determines not just access to schools but also the number of teachers in a school, and this, in turn, significantly affects enrollments. As a consequence, the decision to situate schools in relatively small habitations implies that the benefits of access are less for the residents of these habitations; they are provided with lower quality schools. School location policies, then, contribute to regional schooling inequalities, with children in larger habitations attending better quality schools, as measured by the number of teachers in the school. Further, because the SC/ST habitations are generally smaller than others, the correlation between the number of teachers in a school and habitation size implies that schools in the SC/ST habitations will generally be of lower quality than those located in other habitations.

The simple relationship between habitation size and schooling attainment which derives through a consideration of school access alone is further modified when we recognize that schooling attainment also varies with the size distribution of habitations in a village because of its effect on schooling segregation. Districts with larger habitations are more likely to support segregated schools. The results of the previous section suggest that this improves schooling for the upper castes but not for children from the SCs. These results therefore suggest that the benefits of living in districts with improved access to schools differ across the SC/ST and the upper caste children.

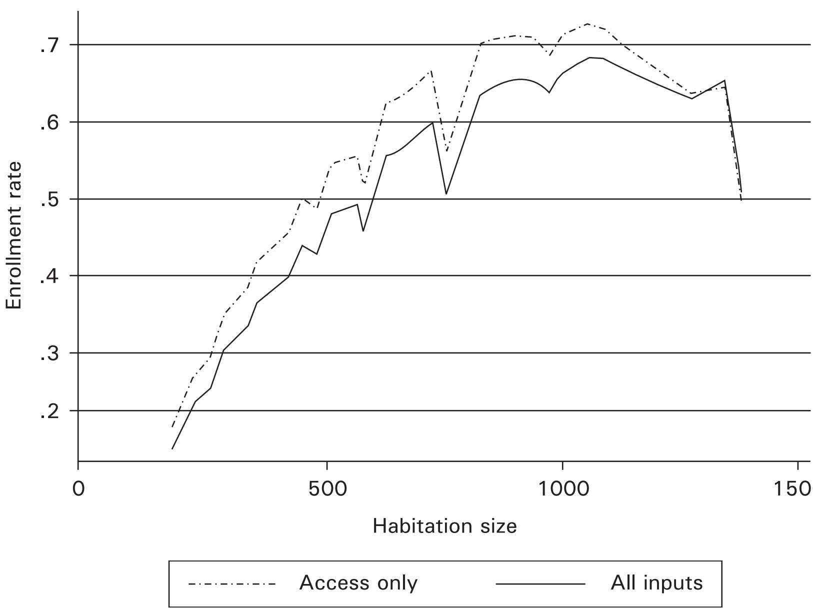
What is the magnitude of these benefits, and how does the auxiliary effect of habitation size on school quality, as measured by teacher strength and schooling segregation, mediate the relationship between habitation size and schooling attainment? The answer to this question can be obtained by combining the first stage regression results on the determinants of school attributes with the IV probit estimates of the effect of school inputs on enrollment. However, because the first stage regressions utilize a number of different measures of habitation size, the overall effect of habitation size on school attributes are difficult to infer from the regression results reported in table 5.

I therefore proceed by first running a reduced form predicting equation for each of the school attributes in question (proportion of habitations with a school located in them, the proportion of schools with two or less teachers, and the probability of a school in a SC/ST habitation) on average habitation size in the district and its square. I use these regressions to predict the effect of average habitation size on the school attribute in question, and then combine this with the IV probit estimates of the effect of school attributes on enrollment to infer how district average habitation size affects enrollments through its effect on school inputs. The IV probit results I use are those reported in table 8, which allow the determinants of schooling enrollment to vary by caste. This correspondingly allows the calculation of the effect of habitation size on schooling attainment to also vary by caste.

I graph the results separately for the SC/ST children and the other children, in figures 8 and 9, respectively. Each figure depicts two graphs. The first explores the effect of habitation size on enrollments, allowing habitation size to affect enrollment only through its effect on school access. The second graph adds on the additional influence of habitation size on school enrollment through its effect on teacher availability and school segregation.

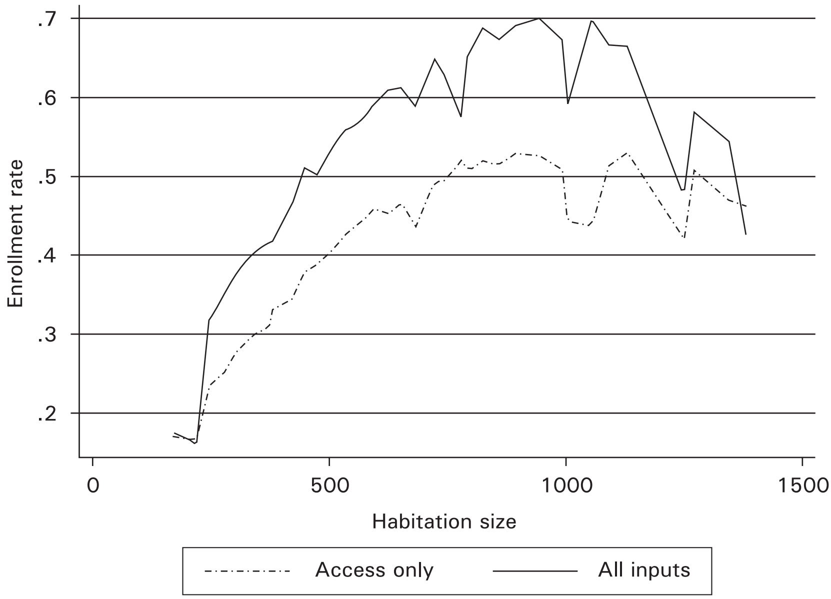
Both graphs confirm the positive effect of the average size of habitations, in a district, on enrollment through its effect on school access: because the probability of a habitation having a school located in it increases with its size, children who live in districts with larger habitations are significantly more likely to be enrolled in school. Calculated at the mean level of the variables, the elasticity of enrollment with respect to habitation size, allowing this effect to operate only through access, is 0.36 for children of the upper castes and as

FIGURE 8. Effect of Average Habitation Size on Enrollment of SC/ST through School Inputs



Source: NSS and AIES various rounds.

FIGURE 9. Effect of Average Habitation Size on Enrollment of Other Castes through School Inputs



Source: NSS and AIES various rounds.

high as 0.48 for those of the SCs/STs. If school location policies determined access to school, without affecting measures of school quality, this suggests that these policies could serve as a powerful means for reducing schooling inequalities—they suggest that providing access to schools has a far larger effect on children from the SCs/STs than on the upper caste children.

This conclusion is modified if we allow habitation size to also affect school quality, as represented by the number of teachers in a school and the extent of schooling segregation. For the upper caste children, the benefits of residing in districts with larger habitations are magnified because schools in these districts have more teachers and are more likely to be characterized by student populations which are segregated by caste. Allowing habitation size to affect enrollment through these additional inputs increases the benefit of a 1 percent increase in habitation size on enrollment from 0.36 to as much as 0.45. However, the benefits of residence in districts with larger habitations do not similarly apply to the SC/ST children. This is primarily because extensive schooling segregation in these districts negatively affects the enrollment decisions of the SC/ST. The elasticity of enrollment with respect to habitation size falls to 0.44, once these auxiliary effects are accounted for.

Thus, despite the fact that the effect of access on the enrollment of the SC/ST substantially exceeds its effect on the upper castes, the overall benefits of residing in districts with widespread access to schools do not significantly differ by caste, suggesting that school location policies cannot be an effective instrument for reducing schooling inequalities. This result is the consequence of the fact that school location policies affect school quality, not just access to schools.

Conclusion

In designing a policy for the construction and location of primary schools in rural India, the government has paid scant attention to the fact that the location of schools determines not just access but also important dimensions of school quality. Because of the geographic distance between habitations in rural India and their small size, the decision to situate a school in a habitation determines total enrollment and hence the total number of teachers in the school. The policy of constructing schools in all habitations above a minimum size in conjunction with caste-based residential segregation also implies that many villages have multiple schools, with the SC/ST children attending different schools from the children of other castes.

Consistent with these hypotheses, this paper documents the significant role of habitation size in determining not just access to schools but also critical school inputs such as the number of teachers. The inability to ensure that each school is of the optimal size, reflected in the effect of habitation size on teacher availability and the consequent variation in teacher strength across schools, implies that school location policies are partly responsible for the significant observed variation in school quality across rural India. Of habitations of size large enough to receive a school, households who reside in larger habitations will be characterized by higher levels of schooling than those who reside in smaller habitations. Because the SC/ST habitations are generally smaller, the government's school location policies imply that schools located in the SC/ST habitations will be of lower quality, thereby contributing to caste-based schooling inequalities.

They also contribute to caste-based schooling inequalities through their effect on school segregation. The results of this paper show that the availability of schools in the SC/ST habitations, which in turn enables schooling segregation, affects schooling attainment. However, the effect of schooling segregation varies by caste; the availability of schools in SC/ST habitations increases the schooling of the other caste children but reduces that of the SC/ST children. Higher levels of schooling segregation in districts with widespread access to schools wipe out the differential benefits of access for the SC/ST.

The results of this paper suggest that improvements in school quality cannot be affected without re-considering school location policies. Improving school quality along the dimensions considered in this paper is, however, not an easy task. The importance of teacher strength suggests that a policy which consolidates habitation schools to provide one school in each village is likely to improve schooling attainment, since it would enable an optimal number of teachers in each school. While the greater distance to school implied by such a consolidation, particularly for children from the SC/ST habitations, may reduce access, the savings generated by the consolidation could be used to implement a system of cash transfers to children from the SCs and the STs, conditional on their school attendance records. Evidence from Mexico (Schultz, 2004) suggests that these transfers can significantly enhance access.

On the other hand, it is unclear what a consolidated school system would imply for schooling segregation. While it could potentially reduce it, the evidence of this paper that schooling attainment is affected by the average quality of the student population suggests a segregating force; in the presence of such externalities, households may be willing to pay to achieve a level of segregation through, for example, their support for private schools (Benabou, 1996). Indeed, one could view the rapid rise of private schooling in rural India as a consequence of the increases in schooling by the SC/ST, and the consequent reduction in the average "quality" of students in government schools. School consolidation, under these conditions, may not significantly alter schooling segregation. Even so, the positive effects on the number of teachers and the improved ability to bear the fixed costs of investments in infrastructure, playgrounds, laboratories, computers, and libraries may still imply that a re-consideration of school location policies may be necessary to ensure improvements in schooling.

Comments and Discussion

Rukmini Banerji: This paper focuses on an interesting and relatively unexplored dimension of schooling provision in India—the effect of school location on both schooling attainment and caste-based segregation in primary schools. Kochar argues that India's school location policies have had a significant impact on school quality. In her paper, school quality is coterminous with school inputs, especially the number of teachers. Using district-level data she shows that school segregation by caste is higher in districts with widespread access to schools, thus the benefits of access are wiped out by the negatives of school segregation as far as school enrollment and school attainment are concerned.

While much of her contribution to the allied literature is empirical, I will raise some conceptual issues to think about.

In the last five years, especially in the north Indian educationally backward states like Uttar Pradesh and Bihar, there has been large-scale recruitment of teachers and para-teachers. For example, in Uttar Pradesh, para-teachers (*shiksha mitras*) are local and recruited by the local village body. Early grades in primary school are generally assigned to the para-teachers. In several states like Uttar Pradesh and Bihar, a significant percentage of para-teachers are female and local. Along with this decentralized recruitment, governments have been building new schools with a special target on unserved habitations. Considerable work has been also been done in upgrading small schools and in constructing additional classrooms. Along with this expansion of school inputs, both at Central and state government levels, there has been a big push to bring the teacher-pupil ratio (TPR) to 1:40.[1]

The data for this paper is based on a variety of sources of data from 1993, 1999, and 2002. More recent data may have allowed Kochar to disentangle some of the new trends in school provision and estimate its implications. For example, without disaggregated data at school level it is difficult to measure the impact of para-teacher (local and perhaps same caste as students

1. ASER stands for Annual Status of Education Report. ASER 2007 sample of 16,000 schools indicates that at the all-India level, the TPR is now below 40. Except for Uttar Pradesh, Bihar, Jharkhand, and West Bengal, many of the major states have TPR below 40 or around 40. See the set of entire reports from 2005–07 on www.pratham.org.

in small habitations) versus regular teacher (higher caste and not local) on schooling attainment. It is widely perceived that para-teachers are less absent and are more effective than regular teachers as they are hired on a contractual basis.

What is the "optimal" school size and for what outcome? Kochar is concerned about schooling attainment. If you consider learning as another outcome, then the available evidence points to surprising and unexplained trends. Bihar—which is a state that has had very poor school provision and woefully inadequate number of teachers until very recently—does well in terms of children's learning in national learning surveys (school-based grade level government administered NCERT pen and paper tests) as well as household-based oral individual assessments of ASER. Tamil Nadu—which is a well-functioning state with adequate school inputs and incentives—does poorly, at least in ASER. Himachal Pradesh has very small schools, (children's enrollment and teacher availability) and a high degree of multi-grade teaching. The conception of, and evidence on, "optimal" school size is worthy of further study.

The rapid increase of private schools is very visible in many states in north India but is poorly measured. Government school statistics do not have a complete enumeration of these schools, especially the un-recognized private schools. Usually, the incidence of private schools is higher in bigger habitations and is lower in the catchment area of a well-run government school. It would be interesting to look at school attainment and school segregation in habitations where there are not only several government schools but also private schools.

While most of India now has primary schools within 1 km of the habitation, understanding implications of school location and provision policies are critical as middle school/upper primary school expansion begins to take place. To date, the government norm for middle schools is that it should be located within 3 km of the habitation. Given this landscape, Kochar's topic is a relevant one and her rigorous empirical research is timely.

Rajnish Mehra

Introduction

I enjoyed reading this thought-provoking paper. It raises an interesting issue regarding the adverse consequences of a well-intentioned government policy. The paper analyzes the effectiveness and consequences of a post-independence priority: the emphasis on facilitating educational achievement for India's rural population, both as a desirable goal in and of itself, and as a crucial input to alleviating entrenched caste and gender educational discrimination. Implementation of this policy focused broadly on *access*—providing a school within walking distance (typically 1 km) of each rural household.

The author undertakes a challenging task—to catalogue and critically assess the consequences of this policy, half a century later—and is to be complimented on her research. Central to understanding the results of this paper is the concept of "habitation": a distinct cluster of houses existing in a compact and contiguous manner[2] with a local name. Further, its population should not be less than twenty-five houses in plain areas and not less than ten in sparsely-populated areas. In case there exist more than one such cluster of houses in a village, they are treated as separate habitations only if the convenient walking distance between them is more than 200 meters. A village may thus have within it one or more habitations.

Basic Findings

Most habitations in India are small and dispersed. Large habitations are characterized by efficient school size, more schools, more teachers per school, and higher rates of enrollment. By contrast, small habitations typically manifest lower enrollment numbers, one or no school with two or less teachers, and generally an inefficient school size. A large village is likely to have habitations segregated by caste and, often consequently, this results in schooling segregation—typified by one small school with two or less teachers and lower enrollment rates.

The paper argues that habitation size may be as, or more important than other attributes such as community income in determining schooling outcomes. There are two implicit assumptions motivating the discussion: 1) Over a certain range, there are increasing returns to scale in school design. This is a reasonable assumption given the large fixed costs of establishing a school and the negligible marginal costs of an additional student. One school with an adequate number of teachers and an integrated student body is better than two inadequately staffed segregated schools. Requiring a school to be located within a specified distance from a household leads to a

2. NCERT's Seventh All India Survey.

sub-optimal size with concomitant effects on quality. 2) There is an optimal location choice that maximizes access while minimizing inefficiencies and externalities.

From a policy point of view, how would this be implemented? To address this we need an objective function and resource constraints. This is a problem in Economic Geography. The paper does not discuss this issue. There are a number of models that can be adapted to address the issues discussed in the paper for example, Marianov and Serra (2004), and Banerji and Fisher (1974). The two key arguments with the maximum weight in such an objective function are attendance by teachers *and* attendance by *students*. Kremer et al. (2004) and others have found that, on average, 25 percent of teachers in rural schools are absent on any given day. This is shown in table 9.

TABLE 9. Teacher Absence in Public Schools by State

State	Absence (%)
Maharashtra	14.6
Kerala	21.2
Haryana	21.7
West Bengal	24.7
Uttar Pradesh	26.3
Punjab	34.4
Bihar	37.8
Jharkhand	41.9
Weighted Average	24.8

Source: Kremer et al. 2004.

They find that absenteeism increases with the distance of the school from the main road. From this point of view, smaller schools local to the habitation might be a critical factor in increasing teacher attendance. The student attendance is also of first-order importance. In light of this, one can understand what motivated the government to opt for the current policy.

The simplest way to increase school participation is to reduce the cost of schooling, or even pay for attendance (Kremer 2003). Mid-day meals have proven to be an effective motivator. This is especially so for the SC/ST students—the elasticity of enrollment with respect to access is probably higher for them. From a welfare point of view, school attendance is desirable even in a segregated system and with few teachers, when the alternative is non-participation (child labor). Discrimination is a ground reality, especially in rural India. Even if schools are located randomly and students are free to choose which school to attend, one suspects that, in equilibrium, segregation may be the outcome (Schelling, 1978).

A recent development is the proliferation of private "unrecognized" schools even in rural areas. These schools are emerging as viable substitutes to government-run schools. Some discussion is warranted on how the presence of these schools can affect educational attainment in districts made worse off by the current government policy.

Sector-specific rates of technological progress are also likely to have important implications for the distributions of schooling and these are likely to be different for landed v. landless households. The results from Foster and Rosenzweig (1996) suggest that technical change in agriculture is likely to increase schooling inequality in rural areas. It increases schooling returns for landed households—who make the decisions about the adoption and management of new seeds but not for landless households—who undertake such tasks as weeding or harvesting crops.

Conclusion

The 'pluralist' view of the causes of poor schooling outcomes in rural India, for example, Drèze and Kingdon (2001), recognizes several key determinants of school participation: household resources, parental motivation, the returns to child labor, and school quality. A valuable insight that emerges from the analysis in this paper is that habitation size may be as or more important than other attributes such as community income in determining schooling outcomes. The next step would be to investigate how one might extract efficiency gains from this insight.

General Discussion

Dilip Mookherjee began the general discussion by echoing the concerns raised by Mehra about the policy alternatives. Because commuting time is a considerable constraint in many situations, there would need to be a way of getting children in small habitations to larger schools in larger habitations. Mookherjee noted how dispersed hamlets in India can be with children often commuting several hours each day. Second, if the schools were desegregated by merger into a larger school, would the lower castes receive decreased attention from teachers and reduced access to resources? Mookherjee argued that large schools would be responsive to concerns about resource costs because they were likely to gain from economies of scale. However, if the primary concern is equality of opportunity for education, resource costs are less important and the current policy may be justified, given all the other constraints.

Abhijit Banerjee made two points. First, rather than interpreting the paper as a statement about optimal resource allocation, it could be seen as a statement about what could be done to generate caste-based equity. Therefore, if schools in small hamlets are doing poorly, it may be better to devote extra resources to those schools rather than thinking of it as a problem of bussing children. Second, he noted that it may be useful to check the exclusion restrictions for the econometric estimations. He argued that it is not clear whether being a small hamlet in a less developed district is the same thing as being a small hamlet in a more developed district.

Another participant suggested that there may be a tradeoff between gender and caste inequality from the perspective of walking further to school. The literature has shown that girl enrollment is more sensitive to distance than that of boys. Second, the caste of the teacher may be important for educational outcomes, as there appear to be big differences in teacher attendance depending on whether a high caste teacher is going to a Scheduled Caste (SC) hamlet or if a SC teacher is going to a SC hamlet.

Indira Rajaraman echoed the gender equity issue raised by others. She noted that one of the well-known reasons why girl children are not sent to school—especially after puberty—is because of the non-availability of toilets in schools. So, the current policy of school access within 1 kilometer has meant that they can go home in between the school day, which has really improved gender equity as far as attendance in schools is concerned.

Finally, several other participants raised additional points. One questioned why the current policy is producing gains to the non-Scheduled Caste groups, and why those gains were asymmetric to the costs to the SC/ST students. Another suggested that if bussing is impractical, other policy responses, such as a cash transfer conditional upon attendance, may be an effective means of achieving the desired result. It was also argued that more than just the location of a hamlet needs to be included in the analysis of performance. The infrastructure and community involvement may also differ greatly among and within hamlets.

The author responded to the questions and comments by first reiterating the main point of the paper that there is huge variation in school quality across schools by habitation size and SC/ST, and that a lot of these caste-based inequalities are a result of the current school location policy. So, in a sense, the paper was meant as an identification of a problem that is not sufficiently recognized in current policy circles.

Regarding policy alternatives, Kochar had several suggestions. First, one could contemplate a different configuration of schools. Rather than the traditional K to 5, and 6, 7, and 8, it could be possible to have schools which go from kindergarten to grade 2, and then moved to a combined school. She remarked that the state of Punjab has actually moved in the past few years to a unified school system. The alternative of providing a conditional cash transfer to SC/ST households only, provided that they go to school would be responsive to Banerjee's comment about the relative ease of targeting the SC/ST.

Finally, Kochar emphasized that many of the points raised during the discussion should be areas for future research. Whether the SC/ST will be treated equally or not when they go to other schools can only be answered with further research. In addition, the distribution of teachers in panchayat schools and the caste relationship of teachers to hamlets are both researchable questions that should be explored in the future.

References

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon. 2001. "Child Health and Social Enrollment: A Longitudinal Analysis." *The Journal of Human Resources* 36 (1): 185–205.

Banerji S. H., and H. B Fisher. 1974. "Hierarchical Location Analysis for Integrated Area Planning in Rural India." *Papers of the Regional Science Association* 33: 177–94.

Benabou, Roland J. M. 1996. "Equity and Efficiency in Human Capital Investment: The Local Connection." *Review of Economic Studies* 63 (2): 237–64.

Birdsall, Nancy. 1985. "Public Inputs and Child Rearing in Brazil." *Journal of Development Economics* 18 (1): 67–86.

Borjas, George J. 1995. "Ethnicity, Neighborhoods, and Human-Capital Externalities." *American Economic Review* 85 (3): 365–90.

Case, Anne, and Lawrence F. Katz. 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth." NBER Working Paper 3705.

Coleman, James. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (Suppl.): S95–S120.

Deolalikar, Anil. 1997. "Determinants of School Enrollment and School Expenditures in Kenya." Unpublished Manuscript, University of Washington, Seattle, WA.

Drèze J., and G. G. Kingdon. 2001. "School Participation in Rural India." *Review of Development Economics* 5 (1): 1–24.

Fernandez, Raquel, and Richard Rogerson. 1996. "Income Distribution, Communities, and the Quality of Public Education." *Quarterly Journal of Economics* 111 (1): 135–64.

Filmer, Deon. 2004. "If You Build It, Will They Come? School Availability and School Enrollment in 21 Poor Countries." World Bank Policy Research Working Paper 3340. Washington, D.C.: The World Bank.

Foster, Andrew D., and Mark R. Rosenzweig. 1996. "Technical Change and Human Capital Returns and Investments: Evidence from the Green Revolution." *American Economic Review* 86 (4): 931–53.

Kochar, Anjini, Kesar Singh, and Sukhwinder Singh. 2007. "Targeting Public Goods to the Poor in a Segregated Economy: An Empirical Analysis of Central Mandates in Rural India." Unpublished Manuscript, Stanford University, Stanford, CA.

Kremer, M. 2003. "Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons." *American Economic Review* 93 (2): 102–106.

Kremer, M., Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer, and F. Halsey Rodgers. 2004. "Teacher Absence In India: A Snapshot." Forthcoming, *Journal of the European Economic Association.*

Lillard, Lee A., and Robert J. Willis. 1994. "Intergenerational Educational Mobility: Effects of Family and State in Malaysia." *Journal of Human Resources* 29 (4): 1126–66.

Marianov, Vladimir, and Daniel Serra. 2004. "New Trends In Public Facility Location Modeling." UPF Economics and Business Working Paper No. 755.

Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic Diversity, Social Sanctions and Public Goods in Kenya." *Journal of Public Economics* 89 (11–12): 2325–68.

Planning Commission India. 1952. *The First Five year Plan*. New Delhi: Planning Commission.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: W.W. Norton and Co.

Schultz, Paul T. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74 (1): 199–250.