School Participation in Rural India*

Jean Drèze
Geeta Gandhi Kingdon

April 2000

Abstract:
This paper presents an analysis of the determinants of school participation in rural north India, based on a recent household survey which includes detailed information on school characteristics. School participation, especially among girls, responds to a wide range of variables, including parental education and motivation, social background, dependency ratios, work opportunities, village development, teacher postings, mid-day meals and infrastructural quality. Mid-day meals are particularly effective: the provision of a mid-day meal in the local school roughly halves the proportion of girls excluded from the schooling system. School quality matters, though it is not related in a simple way to specific inputs.

* Drèze: Centre for Development Economics, Delhi School of Economics, Delhi 110 007, India. Tel: 91-11-725 7005; fax: 91-11-725 7159; e-mail: jean@cdedse.ernet.in. Kingdon: Department of Economics, University of Oxford, Oxford OX1 3UL, United Kingdom. Tel: 44-1865-271065; fax: 44-1865-281447; email: geeta.kingdon@economics.ox.ac.uk. We are grateful to Angus Deaton and Haris Gazdar for valuable suggestions and comments, and to Anuradha De for helping us to master the data set. Jean Drèze's work builds on his involvement with the Network on Inequality and Poverty in Broader Perspective, based at Princeton University. Geeta Gandhi Kingdon's work was supported by the Nuffield Foundation and the Wellcome Trust.

JEL Classification: I21
Abbreviations: none
Number of figures: zero
Number of Tables: 7
Date: 15th February 2000
Contact author: Geeta Kingdon, Department of Economics, University of Oxford, OX1 3UL, United Kingdom. Tel: 44-1865-271065; fax: 44-1865-281447; email: geeta.kingdon@economics.ox.ac.uk.

1. Introduction

About one third of all Indian children are out of school. In the large north Indian states, which account for over 40% of the country's population, the proportion of out-of-school children in the 6-14 age group is as high as 41%, rising to 54% among female children.[1] Considering the crucial role of elementary education in development, the universalisation of schooling in India is one of the most urgent development issues in the world today.

Yet, relatively little is known about the precise reasons why so many Indian children are out of school. In public debates, the tendency is to highlight a single 'explanation'. In official circles, for instance, the problem is often blamed on parental indifference towards education -- a convenient argument since it diverts attention from the responsibility of the state. Others consider that child labour is the overwhelming obstacle: according to the Campaign Against Child Labour (1997), India has more than 60 million child labourers, working 12 hours a day on average. Neither of these single-focus explanations, however, stands up to careful scrutiny (Bhatty *et al*., 1997). This is not to deny that they contain a grain of truth. The real challenge is to build a balanced picture of the determinants of school participation, which integrates different lines of explanations: lack of parental or child motivation, the costs of schooling, the demands of child labour, and the low quality of schooling, among others.

As a modest step in that direction, this paper presents an analysis of the determinants of school participation based on a recent survey of schooling in north India, the PROBE survey.[2] This is not the first study of its kind; earlier analyses of a similar inspiration include Duraisamy and Duraisamy (1991), Duraisamy (1992), Kingdon (1994, 1996, 1998), Jayachandran (1997), Labenne (1997), and Sipahimalani (1998), among others.[3] The PROBE survey, however, offers unusual possibilities for scrutinising different influences on schooling decisions, in so far as it contains detailed information not only on the characteristics of about 4,400 children and their households,

but also on the schools to which these children have access. In particular, this survey enables us to examine the influence of different types of 'school quality' variables on school participation and educational achievements in rural India.

2. The Schooling Situation in Rural India

By way of orientation, we begin with a brief sketch of the schooling situation in rural India, with specific reference to the four major states covered by the PROBE survey: Bihar, Madhya Pradesh, Rajasthan, Uttar Pradesh (hereafter the 'PROBE states').[4] The main features of the schooling situation in these states, as they emerge from the PROBE survey, include the following (see also Table 1).[5]

School availability: Most villages have at least one government primary school (classes 1 to 5). Government schools charge negligible fees, and never refuse to enrol a child. There are no Board examinations until class 5 (in fact, well after class 5 in most states), but primary schools conduct school tests and children are sometimes asked to repeat a particular class. Only a minority of villages -- about 30% -- have a 'middle' school (classes 6 to 8).

Private schools: In addition to government schools, private schools with primary sections are available in a significant minority (about 17%) of villages. Private schools charge fees, and generally attract children from relatively privileged families, though children from poor families are not entirely excluded.

Parental motivation: Qualitative data from the PROBE survey suggest that parental interest in education is generally quite high. Most parents would like their children (particularly sons) to be educated, and favour compulsory education for all children. However, they have a dim view of the schooling system, and endless complaints about teaching standards in government schools.

School participation: In the 6-14 age group, according to the PROBE survey, 85% of boys

and 56% of girls are currently enrolled in school.[6] Among ever-enrolled children in the 13-18 age group, 81% have completed class 5.

Child labour: Time utilisation data from the PROBE survey indicate that out-of-school children work about 4.7 hours a day on average (about 2 hours more than school-going children), mainly helping their parents at home and in the fields. Work hours are a little longer for girls than for boys, and particularly long for eldest daughters in poor families.

School quality: Aside from parental testimonies, the PROBE survey found much direct evidence of the dismal state of government schools. To illustrate: (1) Only one fourth of the sample schools have at least two teachers, two all-weather classrooms, and some teaching aids; (2) If all children aged 6-10 in the sample villages were enrolled in a government primary school, there would be more than 100 pupils per classroom on average, and 68 pupils for each teacher; (3) At the time of the investigators' visit, about half of the schools had no teaching activity.

Silver lining: Aside from the four major states listed earlier, the PROBE survey also covered Himachal Pradesh, a smaller state located in the Himalayan region. In sharp contrast with the other states, Himachal Pradesh had high rates of school participation (96% of all children aged 6-14 were studying), and low educational disparities between boys and girls as well as between different communities. There were also many signs of a higher quality of schooling in Himachal Pradesh, e.g. lower pupil-teacher ratios, better teaching standards, and a more cooperative rapport between parents and teachers. This success is all the more impressive considering that, not so long ago, Himachal Pradesh was widely regarded as a 'backward' region of north India.[7]

3. Issues and Hypotheses

The main focus of this paper is on school participation as a household decision.[8] At a general level, this decision may be thought to depend on the costs and benefits of basic education, broadly

understood. The costs depend both on the opportunity cost of a child's time, and on the direct costs of schooling (e.g. expenditure on fees and books). The benefits include economic returns, mainly in the form of improved earning opportunities as well as more productive work within the household. Other possible benefits include better health, improved social status, greater bargaining power, and the joy of learning, among others. A simple model of school participation decisions in a cost-benefit framework is presented in Drèze and Kingdon (1999), where the limitations of this approach are also discussed. Despite its limitations, the cost-benefit framework provides a useful starting point for interpreting empirical findings on the relation between school participation on the one hand and household, school or village characteristics on the other.

The cost-benefit framework suggests that school participation is positively related to variables that enhance the perceived benefits (or reduce the costs) of education. Income effects, for instance, are likely to be positive. The effect of land ownership, on the other hand, is ambiguous: land is a form of wealth (suggesting a positive effect on school attendance), but it also raises the productivity of child labour within the household, and hence the opportunity cost of school attendance. A similar remark applies to the ownership of farm animals.

Other household characteristics of interest include caste and parental education. It is well known that school participation in India is particularly low among socially disadvantaged communities, notably the 'scheduled castes'. [9] The caste bias, however, requires further scrutiny in at least two respects. First, it is not clear whether (and to what extent) this bias persists after controlling for household income, parental education and related characteristics. Second, an earlier study (Kingdon, 1998) finds that conditional on school enrolment, the achievements of scheduled-caste pupils are no lower than those of other pupils. This can be interpreted as tentative evidence that discrimination within the schooling system is not the root of the problem.[10] The PROBE survey is an opportunity to reexamine this pattern. As far as parental education is concerned, the main issue is whether positive intergenerational effects persist after introducing extensive controls for other

household and school characteristics. If so, social returns to education are even higher than standard estimates suggest. Another issue concerns the respective influences of paternal and maternal education on male and female schooling. One study (Jayachandran, 1997) suggests that intergenerational same-sex effects are stronger than cross-sex effects, i.e. boys' schooling is more responsive to father's education than to mother's and vice-versa for girls. These intergenerational correlations, however, may reflect the influence of missing variables and call for further scrutiny.

Turning to school characteristics, the cost-benefit framework suggests that (1) school quality has a positive effect on initial school participation (i.e. the discrete choice of enrolling a child in class 1), and (2) it also has a positive effect on further investment in education (including the decision to keep a child enrolled in successive classes) if school quality and private expenditure are complementary inputs, which need not be the case but is likely to apply.

Empirical measures of school quality, however, are not easy to devise. In the literature, much attention has been paid to the pupil-teacher ratio (or 'class size'), a controversial school-quality indicator. Evidence of a negative effect of class size on pupil achievements has proved somewhat elusive (see Fuller, 1986, and Hanushek, 1986, 1995), though it does show up in some recent studies based on improved estimation techniques (see e.g. Angrist and Lavy, 1996, and Case and Deaton, 1997). In India specifically, it would surprising if pupil-teacher ratios did not matter. Indeed, overcrowding is commonly mentioned by teachers as one of their major problems, and qualitative observations from the PROBE survey lend much credibility to this concern (The PROBE Team, 1999). So far, however, this issue has not been the object of detailed investigation.[11]

Aside from the teacher-pupil ratio, commonly-used indicators of school quality include teacher salaries, teacher experience or training, expenditure per pupil, and various indicators of physical infrastructure. In the Indian context, however, there is a case for focusing on a different list of school-quality variables. For instance, teacher salaries are unlikely to matter, since salaries bear little relation to qualifications or performance.[12] Similarly, expenditure per pupil has little relevance

in this case, since it is essentially the product of teacher salaries (which account for about 95% of recurrent expenditure) and the pupil-teacher ratio. On the other hand, the qualitative findings of the PROBE survey clearly point to the need to capture other aspects of school quality such as teaching standards, classroom activity and incentive schemes. One goal of this study is to examine the respective influences of different aspects of school quality on school participation. The following indicators were considered, among others: pupil-teacher ratios; physical facilities; the presence of female teachers; teacher attendance rates; frequency of inspection; the provision of mid-day meals; teacher qualifications and training; the frequency and severity of physical punishment; classroom activity levels; and indicators of teacher-parent cooperation.

4. Data and Estimation

The data set

The PROBE survey collected household data in 122 randomly-selected villages of Bihar, Madhya Pradesh, Rajasthan, Uttar Pradesh and Himachal Pradesh. In each village, all school facilities were surveyed and a random sample of 12 households were interviewed. Further details of the sampling procedure and related matters are given in Drèze and Kingdon (1999).

Our primary objective is to identify the determinants of school attendance and educational attainment. Specifically, we focus on the following dependent variables (in each case, individual children are the basic units of observation): (1) Initial enrolment: This is a dummy variable taking value 1 if the child has ever been enrolled in a school, and 0 otherwise. (2) Current enrolment: A dummy variable taking value 1 if the child is currently enrolled, and 0 otherwise.[13] (3) Grade attainment: This is the highest grade achieved by the child.

Our interest is in primary schooling (the main focus of the PROBE survey itself). Accordingly, when the left-hand side variable is 'current enrolment', the observations are restricted

to children in the 5-12 age group.[14] When 'grade attainment' is the dependent variable, the reference group consists of children aged 13-18 (i.e. children who are supposed to have completed primary schooling). For 'initial enrolment', children aged 5-18 are taken as the reference group.

The right-hand side variables consist of individual characteristics, household characteristics, school characteristics and village characteristics. They are listed in Table 2a, together with their sample means. Precise definitions, and some explanatory notes, are given in Table 2b. We proceed with further comments on specific variables.

Teacher inputs

As discussed earlier, the relation between pupil-teacher ratios and pupil achievements has received sustained attention in the literature. Earlier studies typically have test scores (or some related measure of pupil achievements) on the left-hand side, and the pupil-teacher ratio (or some instrument for it) on the right-hand side. In the present case, however, the left-hand side variable is an indicator of school participation such as 'initial enrolment'. This makes it inappropriate to put the pupil-teacher ratio on the right-hand side, since the latter is directly affected by enrolment rates.

One step around this endogeneity problem is to focus on the child-teacher ratio in the village (*i.e.* the total number of children divided by the total number of teachers at the primary level) as an alternative indicator of teacher inputs.[15] This is a useful indicator of teacher inputs in its own right. The PROBE survey suggests, for instance, that school participation is influenced by 'enrolment drives' undertaken by teachers at the beginning of the year; the effectiveness of these enrolment drives is likely to be sensitive to the child-teacher ratio in the village. This is not the end of the endogeneity problem, however, since the number of teachers posted in a village may be sensitive to school participation rates. Indeed, teacher postings are supposed to be geared partly to an official norm of 40:1 for the pupil-teacher ratio.

The official norm, however, is routinely violated. For one thing, it is undermined by an

overall shortage of teachers. Further, teachers tend to lobby for convenient postings (e.g. in more accessible villages), and their pressures are quite effective, leading to a highly uneven distribution of teachers across villages. In the extreme case where teacher postings are entirely determined by exogenous village characteristics, the endogeneity problem disappears.[16]

In this paper, we instrument for the child-teacher ratio (CTRATIO) using 'village size' and 'distance from the nearest road' as identifying instruments.[17] Other treatments of the child-teacher ratio were explored in Drèze and Kingdon (1999). As shown there, the coefficients of other variables are quite robust with respect to different treatments of CTRATIO. Further, different approaches point to similar results as far as the effect of the child-teacher ratio itself is concerned.

School quality indicators

As shown in Table 2, the regressions include six indicators of school quality (other than the child-teacher ratio): whether the school provides a mid-day meal (LUNCH); an indicator of infrastructure (CLASSROOMS); whether the school building is water-proof (WATERPROOF); whether the school was open for 6 days during the 7 days preceding the survey (DAYS6); the number of days spent by the class-1 teacher in non-teaching duties during the 28 days preceding the survey (NTDUTIES);[18] and an index of parent-teacher cooperation (PTCOOP).

One limitation of some of these indicators (particularly DAYS6 and NTDUTIES) is that they may reflect transient circumstances rather than durable characteristics of the school. For instance, DAYS6 is an indicator of school activity during the week preceding the survey, and that too without adjustment for disruptions unrelated to the quality of the school. This limitation is particularly serious when the left-hand side variable is 'grade attainment', which pertains to children (currently aged 13-18) who went through primary school several years before the survey.

Aside from those listed in Table 2, we tried a number of other school-quality indicators, e.g. whether the school conducts regular tests, the proportion of female teachers, the frequency of

physical punishment, whether any teacher was absent on the day of the survey, a dummy for the presence of a private school in the village, and various characteristics (such as training and qualifications) of the class-1 teacher. Though they usually had the 'right' sign, these variables had unstable coefficients, and were seldom statistically significant. To reduce multicollinearity problems, we dropped them from the regressions reported in the next section. In some cases at least, the poor performance of these variables is likely to be related to the limitations mentioned in the preceding paragraph. For instance, while teacher regularity is likely to be an important aspect of school quality, available information on teacher absenteeism pertained only to the day of the survey, and turned out to be too 'noisy' to yield useful results.

The variable PTCOOP is a tentative index of parent-teacher cooperation, calculated as an unweighted average of four dummy variables. The latter take value 1, respectively, if (1) the headmaster reported having approached the parents for help during the twelve months preceding the survey, (2) the headmaster is satisfied with parents' responses to his or her demands, (3) the school has a parent-teacher association, (4) the headmaster considers the parents' attitude towards the school as 'helpful'. The problem of transient influences obviously applies in this case, and it is interesting that in spite of that the parent-teacher cooperation variable has a statistically significant, positive influence on grade attainment. On the whole, the school-quality indicators that turned out to be significant are those pertaining to relatively durable characteristics of the schools.

Aggregation over schools

A majority (77%) of the sample villages have a single school. For villages with more than one school (usually two), the school-quality indicators listed in Table 2 were averaged across schools within a village, using the numbers of children enrolled in each school as weights. The value of an aggregated school-quality indicator calculated in this way may be interpreted as the level of school-quality which a child can expect to get, assuming that he or she is assigned at random

between the different schools in the village, and that the probability of joining a particular school is proportional to current enrolment.[19]

Parental motivation

Our regression variables include one indicator of parental motivation (MOTIV). This is a dummy variable taking value 1 if the main respondent answered 'yes' to the question 'is it important for a girl to be educated?', and 0 otherwise.[20] It may be objected that negative answers to this question could reflect ex-post rationalization on the part of parents, rather than actual motivation: some parents might report that it is not important for their daughter to go to school simply to reconcile themselves to the fact that she is unable to go for other reasons. Note, however, that the question refers to 'a girl', not 'your daughter'. This phrasing reduces the problem of ex-post rationalization, without perhaps eliminating it entirely. It is also worth noting that MOTIV turns out to be a strong predictor of school participation for boys as well as for girls (see below).

Estimation

Since the dependent variables 'initial enrolment' and 'current enrolment' are both binary variables, the discrete-choice probit or logit model is a natural estimation framework. We employ the familiar binary logit, based on maximum likelihood methods. In the case of grade attainment, however, there is no obvious estimation procedure. If never-enrolled children are discarded, OLS estimates of grade attainment are vulnerable to selection bias. We considered a Heckman selectivity-correction model for grade attainment conditional on enrolment, but rejected this approach in the absence of credible exclusion restrictions (the latter involve identifying variables that affect the probability of enrolment but not grade attainment). As an alternative, we decided to retain the never-enrolled children and use an 'ordered logit' model, where grade attainment is reorganised into three hierarchical categories. The categorical grade-attainment variable takes value

0 for never-enrolled children, 1 for ever-enrolled children who have not completed the primary stage (i.e. five years of schooling), and 2 for those who have completed the primary stage.

5. Main Findings: school attendance

The basic results are presented in Tables 3, 4, and 5.[21] Each table focuses on a different left-hand side variable: 'initial enrolment', 'current enrolment', and 'grade attainment' respectively.

We present separate regressions for boys and girls. The 'girls' regressions have a larger number of significant coefficients than the 'boys' regressions. This makes sense, since there is far more variation on the left-hand side in the case of girls. For the same reason, the results obtained when boys and girls are pooled are quite similar to those applying to 'girls only'. For instance, in Tables 3 and 4, all the variables that are statistically significant in the pooled regression are also significant in the 'girls only' regression.[22] In the pooled regressions, the gender dummy (male=1) is always positive and highly significant, indicating a sharp gender bias in school participation.

The contrast between the 'girls' and 'boys' regressions carries another important message: female school participation is not just a 'by-product' of male school participation.[23] Even for given patterns of male participation, female participation varies a great deal depending on household, school and village characteristics. Further, many of these characteristics are responsive to public policy, pointing to the possibility of reducing the gender bias in school participation.

A number of useful insights arise from the regressions presented in Tables 3 to 5. For expositional clarity, the remainder of this section focuses on the 'initial enrolment' and 'current enrolment' regressions (Tables 3 and 4). These two schooling outcomes will be jointly referred to as 'participation'. 'Grade attainment' results are discussed in the next section. Looking across Tables 3 and 4 in a given column, we find a high degree of consistency between the results pertaining to the two different left-hand side variables. This is reassuring: if the results were spurious, we would not

expect them to be robust to the choice of age group and school-participation indicator.

Household variables

The 'household' variables tend to perform better than the 'school' or 'village' variables, not surprisingly since the household variables are more versatile indicators of the circumstances of a child. The household variables almost always have the expected sign and are often statistically significant (especially for girls). Variations of the baseline regressions indicate that their coefficients are quite robust. 'Marginal effects' indicate how a given variable affects the probability of enrolment. To illustrate, the marginal effect of the gender dummy in Table 4 indicates that the probability of current enrolment is about 14 percentage points higher for a boy than for a girl, all other variables being held constant.[24]

As expected, the probability of school participation increases with parental education (both maternal and paternal), though mother's education does not have a significant effect on male school participation. In that sense, inter-generational 'cross-sex' effects are weaker than inter-generational 'same-sex' effects, much as in Jayachandran (1997). The largest inter-generational effect is that of maternal education on girls' school participation.[25] Also as expected, household wealth (as captured by ASSET) enhances school participation for boys as well as girls, and the effect is highly significant for girls. However, consistent with our earlier discussion, land ownership (LANDOWN) has a negative sign for girls (not significant), and similarly with ownership of domestic animals (COWGOAT). The latter has a significant negative impact on girls' current enrolment. So does the dependency ratio (DEPEND), as one would expect not only because the latter is correlated with poverty but also because eldest daughters in households with many children are often expected to look after younger siblings at home (The PROBE Team, 1999, pp. 28-31).

Even after controlling for other household variables, children belonging to 'scheduled castes and scheduled tribes' (SC/ST) and 'other backward castes' (OBC) are less likely to go to school

than children belonging to the general castes (default category). This effect is particularly strong for girls. The coefficient on the MUSLIM dummy is negative, but not statistically significant. Interestingly, this applies even if MOTIV is dropped from the regressions. This goes against the common notion that Muslim culture is inimical to schooling. The fact that school participation is lower among Muslims seems to have more to do with tangible disadvantages such as poverty and low levels of parental education.

Our indicator of parental motivation (MOTIV) is highly significant in all the regressions presented in Tables 3 and 4. The chance of a girl being currently enrolled rises by as much as 30 percentage points if her parents consider that education is 'important' for female children. Interestingly, the chance of a boy being currently enrolled also rises significantly (by 10 percentage points) with this motivation dummy. Even after allowing for an element of spuriousness here (e.g. due to the 'ex post rationalization' factor), the influence of parental motivation seems to be quite strong in comparison with that of most other variables.

It is also worth noting that, when MOTIV is excluded from the regression, the coefficients of CASLAB, SCST, OBC and MUSLIM become larger, and have larger t-ratios. For instance, CASLAB has a significant negative coefficient in the 'current enrolment' regressions when MOTIV is excluded. In other words, the overall educational disadvantage of children belonging to underprivileged social groups is partly mediated by lower parental motivation. The fact that a disadvantage remains (especially for SC/ST children) even after controlling for parental motivation suggests that social discrimination in the schooling system may also be involved. The latter hypothesis is consistent with other findings of the PROBE survey (The PROBE Team, 1999, pp. 49-51). The apparent persistence of an overall bias against SC/ST children in the schooling system is all the more striking considering that pupil incentives are often targeted in their favour.

School variables

The regressions in Tables 3 and 4 suggest that school variables have relatively little influence on primary-school participation among boys. However, this influence is bound to be hard to capture, partly due to the low variation in male school participation in this data set, and partly due to the inherent limitations of our 'school quality' indicators.

None of the school-quality variables are significant in the 'current enrolment' regression for boys. In the 'initial enrolment' regression (again for boys), non-teaching duties (NTDUTIES) and the child-teacher ratio (CTRATIO) are significant, with a negative sign as expected. The provision of a mid-day meal (LUNCH) is also statistically significant, but with a puzzling negative sign. One possible explanation is that school meals are targeted at disadvantaged areas -- see below.

The school variables perform better in the corresponding regressions fo girls. In both regressions ('initial enrolment' and 'current enrolment'), the four variables that capture relatively durable school characteristics (LUNCH, CTRATIO, WATERPROOF and CLASSROOMS) have the expected sign; they are also statistically significant, except for CLASSROOMS in the 'current enrolment' regression. Variables reflecting more transient school characteristics (DAYS6, NTDUTIES and PTCOOP) have the expected sign, but are not significant in most cases.[26] However, there is some evidence of a positive impact of teacher regularity, as captured by DAYS6.

Female school participation is about 15 percentage points higher when the local school provides a mid-day meal than when it does not (Table 4). This roughly corresponds to reducing the probability of non-participation by half - a commendable achievement. A similar result holds for 'initial enrolment' (Table 3). The fact that LUNCH has a significant positive coefficient and a large marginal effect in both female enrolment equations is of some practical importance, and we submitted it to further scrutiny. This finding turns out to survive alternative specifications of these equations (and it shows up again in the 'grade attainment' equation – see below). Nor is it likely to reflect the fact that school meals are targeted at villages that have favourable unobserved characteristics. In fact, the PROBE survey suggests that school meals are more likely to be targeted

at disadvantaged areas. Consistent with this hypothesis, a regression of LUNCH on village characteristics yields a negative coefficient for VDEVELOP (the village development index), and a positive coefficient for 'distance from the nearest road'. The coefficients, however, are not statistically significant. The only statistically significant variable is VVEC (presence of a 'village education committee'), with a positive coefficient.

In Tables 3 and 4, the marginal effect of WATERPROOF (a dummy indicating whether the school building is water-proof) is almost as large as that of LUNCH. Water-proofness is important both because leaking roofs cause prolonged disruptions of school activity during the monsoon, and also as an indicator of infrastructural quality. Based on the latter interpretation, the size of the marginal effect of WATERPROOF is plausible.[27] The coefficient of CTRATIO, for its part, is relatively small: it implies that, say, doubling teacher postings would raise girls' current enrolment by about 3.5 percentage points (as a linear approximation). However, this may be an underestimate of the true effect of teacher inputs, because having more teachers probably enhances other school-quality indicators as well. It is plausible, for instance, that well-staffed schools are more likely to have a proper roof, a mid-day meal, and a high level of parent-teacher cooperation. Even parental motivation may be positively influenced by the number of teachers in the village.

Village variables

As with the school variables, the village variables (VVEC, VDEVELOP and VWASSOC) have little influence on boys' school participation. Turning to girls' participation, the village development index (VDEVELOP) has a positive and statistically significant coefficient in both regressions. One possible interpretation is that the returns to education are higher in better-developed villages. However, the absence of a corresponding effect for boys (even in the 'grade attainment' equation below) reduces the plausibility of this interpretation. Alternatively, VDEVELOP may act as a proxy for community effects: in better-developed villages, school

participation in neighbouring households is likely to be higher, and this, in turn, may enhance school participation in the observed household.[28] The fact that community effects turn out to be much stronger for girls than for boys is entirely plausible, given that the social dimension of school participation is particularly important in the case of girls (The PROBE Team, 1999, pp. 23-25).

The dummy indicating whether the village has a women's association (VWASSOC) also has a positive coefficient, and is significant in the 'current enrolment' regression. However, some reverse causation may be involved here, if women's associations tend to spring up in villages with relatively high levels of female education. Finally, the presence of a village education committee (VEC) appears to have no significant impact on school participation.[29] This is consistent with the notion that these committees are, in most cases, token institutions (The PROBE Team, 1999). It is also possible, however, that the positive effect of village education committees on school participation is entirely mediated by variables such as LUNCH and MOTIV. As noted earlier, school meals are more common in villages that have a village education committee.

6. Main Findings: Grade Attainment

The results for 'grade attainment' (Table 5) are largely similar to - and consistent with - those discussed in the preceding section for 'initial enrolment' and 'current enrolment'. For instance, much as before we find that (1) parental education matters, especially for girls, with the largest marginal effects pertaining to the influence of maternal education on girls' grade attainment; (2) high dependency ratios have an adverse effect on schooling; (3) mid-day meals, village development and the presence of a women's association have a positive effect on girls' attainments; and (4) pupil attainments respond to the child-teacher ratio and infrastructural quality (as captured by WATERPROOF). A few specific points are worth noting.

Social disadvantage: As noted earlier, children from SC/ST or OBC families have relatively low chances of being enrolled. The evidence on grade attainment is less clear-cut: in Table 5 (first column), the coefficient of SC/ST is negative but not significant, and that of OBC is not even negative. In the case of SC/ST children, the results in Table 5 are best regarded as consistent with those of Tables 3 and 4, in so far as the drop in t-value can be attributed to a smaller sample size. In the case of OBC children, however, Table 5 qualifies the earlier results.

Parent-teacher cooperation: Our index of parent-teacher cooperation, PTCOOP, has a positive and significant effect on grade attainment for boys as well as girls. This finding is particularly interesting in light of the rudimentary nature of this index, and is highly consistent with qualitative observations from the PROBE survey. 'School quality' may have far more to do with this kind of intangible input than with standard quantitative indicators such as teacher salaries, physical infrastructure or class size. Unfortunately, measuring the subtler aspects of school quality is a challenging task, only partly achieved by the PROBE survey.

Mid-day meals: As with 'initial enrolment' and 'current enrolment', mid-day meals have a major positive effect on girls' grade attainment. The chances of completing primary education are

30 percentage points higher for girls living in villages with a mid-day meal than for other girls. As discussed in section 5, this does not seem to be due to the fact that mid-day meals are targeted at privileged villages. A more plausible explanation is simply that mid-day meals drive down the private costs of schooling. As observed by The PROBE Team (1999, p.97): '... parents are not generally opposed to female education, but they are reluctant to pay for it. School meals could make a big difference here, by reducing the private costs of schooling.'[30]

7. Himachal Pradesh

As mentioned earlier, Himachal Pradesh has remarkably high levels of school participation by north Indian standards. On the other hand, if we re-run the baseline regressions after adding a dummy variable for Himachal Pradesh, this variable is not statistically significant. This suggests that the high rates of school participation in Himachal Pradesh are fully 'explained' by the household, school and village characteristics included in the regressions.

To shed further light on these issues, Table 6 compares the means of the regression variables for Himachal Pradesh with the corresponding means for the whole sample. For each variable, the table also indicates how the difference between the Himachal Pradesh mean and the overall sample mean would shift the predicted value of 'current enrolment' for girls, based on the marginal effects estimated in Table 4.[31] For instance, the table indicates that the higher level of father's education in Himachal Pradesh raises the chance of a girl being currently enrolled by about 3 percentage points from the sample mean. If we add up all these 'shift effects', we find that the predicted level of female school participation in Himachal Pradesh is about 31 percentage points above the overall sample mean, which is fairly accurate (see Table 1). About half of the aggregate shift effect is accounted for by household variables, and the other half reflects village and school variables. In the latter group of variables, the largest positive shifts are contributed by the village development index

(VDEVELOP), the dummy for women's associations (VWASSOC), and the infrastructural quality indicator (WATERPROOF). The parent-teacher cooperation index (PTCOOP) is also much higher in Himachal Pradesh than elsewhere, and while this has relatively little effect on current enrolment, it does play a significant part in raising 'grade attainment'.

It is worth noting that each of these four variables has something to do with community participation and the social dimension of education. We have already commented on the possible interpretation of VDEVELOP as a proxy for community effects. Women's associations and the parent-cooperation index capture two aspects of community participation. Even WATERPROOF relates to community participation: in Himachal Pradesh, village communities are often involved in the maintenance or improvement of school buildings (The PROBE Team, 1999, p. 124). Seen in this light, our findings support the notion that community participation and civic cooperation have played a major role in Himachal's success, as argued at greater length in the PROBE report.

8. Concluding Remarks

To conclude, a number of valuable insights emerge from this analysis. First, the results lend support to a 'pluralist' view of the causes of educational deprivation in rural India, which gives due recognition to several key determinants of school participation: household resources, parental motivation, the returns to child labour, and school quality. We find evidence of each of these influences, as both common sense and elementary analysis would predict.

Second, we find strong inter-generational effects (i.e. children of educated parents are more likely to go to school), even after controlling for a wide range of variables. Boys' schooling is more responsive to father's education than to mother's, and vice-versa for girls. Maternal education has a large positive effect on a daughter's chances of completing primary school.

Third, scheduled-caste children have an 'intrinsic disadvantage', in the sense of a relatively low chance of going to school even after controlling for household wealth, parental education and motivation, school quality, and related variables. This suggests the persistence of an overall bias against scheduled-caste children in the schooling system, in spite of positive discrimination in pupil incentives. We found no evidence of intrinsic disadvantage among Muslim children.

Fourth**,** the high level of school participation in Himachal Pradesh is entirely accounted for by the variables included in this analysis. About half of the difference in participation rates between Himachal Pradesh and other states can be attributed to household characteristics, with the other half reflecting village and school characteristics. The results also suggest that community effects have played a major role in Himachal Pradesh's success, as argued in The PROBE Team (1999).

Fifth, school meals have a major positive effect on female school participation. The provision of a mid-day meal in the local school roughly halves the proportion of girls excluded from the schooling system. This finding is consistent with the perceptions of parents and teachers (The PROBE Team, p.95), and strengthens the case for school meal programmes (Drèze, 1998).

Finally, the results suggest that school participation and/or grade attainment are positively influenced by several school-quality variables other than mid-day meals, including infrastructural quality, teacher regularity, parent-teacher cooperation, and the number of teachers per child. However, considerable measurement problems arise in capturing school quality. Infrastrucural facilities or even teacher inputs may not matter as much as the functioning of the schools, which is difficult to observe. While the PROBE survey went further than most in that respect, much scope for improvement remains, both in terms of refining school-quality indicators and in terms of combining quantitative analysis with other approaches. The quest for reliable evidence on the relationship between school participation and different aspects of school quality continues.

References

Angrist, J.D., and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114 (1999):533-75.

Bhatty, Kiran, "Educational Deprivation in India: A Survey of Field Investigations," *Economic and Political Weekly*, 4 July and 10 July (1998).

Bhatty, K., De, A., Jean Drèze, A.K. Shiva Kumar, A. Mahajan, C. Noronha, Pushpendra, A. Rampal, and M. Samson, "Class Struggle," *India Today* 13 October (1997):68-73.

Case, Anne, and Angus Deaton, "School Inputs and Educational Outcomes in South Africa," *Quarterly Journal of Economics* 114 (1999):1047-84.

Campaign Against Child Labour, "Public Hearing on Child Labour, Reference Kit," document prepared for the 2nd National Convention on Child Labourers, New Delhi, 1997.

Deaton, Angus, *The Analysis of Household Surveys*, Baltimore: Johns Hopkins, 1997.

Drèze, Jean, "Primary Priorities: Managing Meals," *Times of India*, April (1998).

Drèze, Jean, and Geeta Gandhi Kingdon, "School Participation in Rural India," Discussion Paper No. 18, Development Economics Discussion Programme, STICERD, London School of Economics, 1999.

Drèze, Jean, and Naresh Sharma, "Palanpur: Population, Economy, Society," in Lanjouw, P., and Stern, N.H. (eds.), *Economic Development in Palanpur over Five Decades*, Delhi and Oxford: Oxford University Press, 1998.

Duraisamy, P., "Gender, Intrafamily Allocation of Resources and Child Schooling in South India," Discussion Paper No. 667, Economic Growth Center, Yale University, 1992.

Duraisamy, P., and M. Duraisamy, "Impact of Public Programs on Fertility and Gender Specific

Investment in Human Capital of Children in Rural India," in Schultz, T.P. (ed.), *Research in Population Economics*, 7, 1991.

Fuller, Bruce, "Raising School Quality in Developing Countries: What Investments Boost Learning?," World Bank Discussion Paper No. 2, Washington D.C., 1986.

Glewwe, Paul, and Hanan Jacoby, "Student Achievement and Schooling Choice in Low Income Countries: Evidence from Ghana," *Journal of Human Resources* 29 (1994):843-64.

Hanushek, Eric, "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24 (1986):1141-77.

Hanushek, Eric, "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer* 10 (1995):227-46.

Heyneman, S., and W. Loxley, "Influences on Academic Achievement Across High and Low Income Countries: A Re-Analysis of IEA data," *Sociology of Education* 55 (1982):13-20.

Heyneman, S., and W. Loxley, "The Effect of Primary School Quality on Academic Achievement Across Twenty Nine High and Low Income Countries," *American Journal of Sociology* 88 (1983):1162-92.

International Institute for Population Sciences, *National Family Health Survey 1992-3* Bombay: IIPS, 1995.

Jayachandran, Usha, "The Determinants of Primary Education in India," M.Sc. thesis, Department of Economics, Delhi School of Economics, 1997.

Jensen, Robert T., "Patterns, Causes and Consequences of Child Labour in Pakistan," mimeo, Center for International Development, Harvard University, 1999.

Kingdon, Geeta Gandhi, "An Economic Evaluation of School Management-Types in Urban India: A Case Study of Uttar Pradesh," D.Phil. thesis, University of Oxford, 1994.

Kingdon, Geeta Gandhi, "Student Achievement and Teacher Pay," Discussion Paper 76, Development Economics Research Programme, STICERD, London School of Economics, 1996.

Kingdon, Geeta Gandhi, "Education of Females in India: Determinants and Economic Consequences: A Case Study of Urban Uttar Pradesh," mimeo, McNamara Fellowships Office, Economic Development Institute of the World Bank, Washington DC, 1998.

Labenne, Sophie, "The Determinants of Child Labour in India," mimeo, Department of Economics, Université de Namur, Belgium, 1997.

Moulton, Brent, "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables in Micro Units," *Review of Economics and Statistics* 72 (1990):334-38.

National Council of Applied Economic Research, *Human Development Profile of India: Inter State and Inter Group Differentials. Volume I: Main Report*, New Delhi: NCAER, 1996a.

National Council of Applied Economic Research, *Human Development Profile of India: Inter State and Inter Group Differentials. Volume II: Statistical Tables,* New Delhi: NCAER, 1996b.

Rao, Vijayendra, "Wife-Abuse, its Causes and its Impact on Intra-Household Resource Allocation in Rural Karnataka," in Krishnaraj, M., Sudarshan, R., and Shariff, A. (eds.), *Gender, Population and Development,* Delhi: Oxford University Press, 1998.

The PROBE Team, *Public Report on Basic Education in India,* New Delhi: Oxford University Press, 1999.

Sipahimalani, Vandana, "Education in the Rural Indian Household: A Gender Based Perspective," mimeo, Department of Economics, Yale University, 1997.

Table 1
Basic statistics on sample villages and households

	PROBE states (Bihar, Madhya Pradesh, Rajasthan, Uttar Pradesh)		**Himachal Pradesh**	
Number of sample villages	188		48	
Number of sample households	1221		154	
	girls	**boys**	**girls**	**boys**
Number of children 6-14 in sample households	1362*	1558	166	163
Proportion of children 6-14 in sample households:				
Enrolled in a school	56.2*	85.4	94.6	97.5
Not enrolled	43.8*	14.6	5.4	2.5
Total	100.0	100.0	100.0	100.0
	government	**private**	**government**	**private**
Number of sample schools	195	41	48	6
Proportion of sample schools that are:				
Primary	83.1	61.0	93.7	66.7
Middle with a primary section	16.4	29.3	2.1	33.3
Secondary with a primary section	0.5	9.7	4.2	0.0
Total	100.0	100.0	100.0	100.0
	female	**male**	**female**	**male**
Number of teachers in primary sections	161	675	90	102
Proportion of teachers in primary sections of:				
Government schools	85.1	76.0	75.6	97.1
Private schools	14.9	24.0	24.4	2.9
Total	100.0	100.0	100.0	100.0

* These figures involve *ad hoc* corrections for under-enumeration of adolescent girls; our own analysis does not include such corrections.
Source: The PROBE Team (1999, p.7). The information given in this table pertains to the full set of 188 villages covered by the PROBE survey. The analysis reported in this paper, however, is based on a sub-set of 122 villages where household data were collected (see Appendix for details).

Table 2a
Mean values of regression variables

Variable	Children 5-18 years old (pertaining to Table 3)			Children 5-12 years old (pertaining to Table 4)			Children 13-18 years old (pertaining to Table 5)		
	All	**Boys**	**Girls**	**All**	**Boys**	**Girls**	**All**	**Boys**	**Girls**
Individual characteristics									
MALE	0.559	1.000	0.000	0.568	1.000	0.000	0.528	1.000	0.000
AGE	9.841	9.842	9.841						
AGESQ	109.368	109.109	109.696						
AGE5				0.122	0.107	0.142			
AGE7				0.162	0.162	0.163			
AGE8				0.154	0.150	0.159			
AGE9				0.114	0.121	0.105			
AGE10				0.119	0.132	0.103			
AGE11				0.093	0.099	0.084			
AGE12				0.104	0.099	0.111			
AGE14							0.223	0.198	0.251
AGE15							0.165	0.158	0.173
AGE16							0.166	0.163	0.170
AGE17&18							0.215	0.255	0.170
Household characteristics									
EDU_MO	1.278	1.125	1.472	1.341	1.203	1.522	1.074	0.853	1.321
EDU_FA	5.172	4.921	5.491	5.320	5.115	5.590	4.689	4.238	5.193
CASLAB	0.156	0.165	0.145	0.162	0.171	0.151	0.137	0.145	0.128
JOB	0.103	0.090	0.119	0.095	0.081	0.114	0.127	0.123	0.131
ASSET	10.052	9.433	10.836	9.470	8.811	10.336	11.951	11.618	12.324
PCCRMS	1.343	1.231	1.486	1.308	1.224	1.417	1.459	1.254	1.688
COWGOAT	3.473	3.501	3.438	3.449	3.458	3.437	3.550	3.650	3.438
LANDOWN	3.674	3.522	3.865	3.595	3.448	3.789	3.929	3.783	4.093
DEPEND	1.449	1.389	1.524	1.436	1.384	1.506	1.488	1.408	1.578
MUSLIM	0.098	0.105	0.090	0.100	0.106	0.093	0.091	0.100	0.081
SCST	0.316	0.325	0.306	0.323	0.330	0.314	0.294	0.305	0.282
OBC	0.298	0.291	0.307	0.293	0.284	0.306	0.315	0.318	0.313
MOTIV	0.894	0.878	0.915	0.896	0.880	0.918	0.887	0.870	0.905
Community characteristics									
VVEC	0.557	0.539	0.580	0.545	0.527	0.570	0.595	0.583	0.609
VDEVELOP	1.396	1.304	1.514	1.378	1.298	1.483	1.456	1.323	1.606
VWASSOC	0.139	0.121	0.161	0.135	0.120	0.156	0.150	0.125	0.179
School characteristics									
LUNCH	0.084	0.081	0.086	0.080	0.075	0.086	0.096	0.105	0.087
CLASSROOMS	1.204	1.135	1.291	1.187	1.147	1.239	1.258	1.090	1.445
WATERPROOF	0.338	0.335	0.340	0.352	0.354	0.351	0.289	0.270	0.309
DAYS6	0.350	0.348	0.352	0.349	0.345	0.354	0.352	0.357	0.346
NTDUTIES	1.628	1.452	1.852	1.595	1.351	1.916	1.737	1.805	1.662
PTCOOP	1.379	1.358	1.405	1.378	1.363	1.398	1.382	1.341	1.427
CTRATIO	71.637	74.611	67.870	72.243	74.779	68.904	69.661	74.021	64.789

Table 2b
Description of variables

Variable	Description
Individual and household characteristics	
MALE	Dummy variable: 1 for boys, 0 for girls.
AGE5 to AGE17	Age dummies: 1 for children of relevant age (e.g. five for AGE5), 0 otherwise.
AGE_CH	Child's age in years.
AGESQ	Square of child's age in year.
EDU_MO	Years of education of mother.
EDU_FA	Years of education of father.
CASLAB	Dummy: 1 if the household's main occupation is casual labour, 0 otherwise.
JOB	Dummy: 1 if household's main occupation is regular wage employment, 0 otherwise.
ASSET	Index of assets owned by the household constructed as follows from owned assets: asset = (2*number of watches) + (5*number of cycles) + (2*number of radios)+ (7*number of televisions)+(50*number of motorbikes).
PCCRMS	Number of *pucca* rooms in the house.
COWGOAT	Total number of cows, buffaloes, and goats owned by the household.
LANDOWN	Amount of land owned by the household, in acres.
DEPEND	Dependency ratio in the household: number of children (age 0-18) divided by number of adults.
MUSLIM	Dummy: 1 for Muslim households, 0 otherwise.
SCST	Dummy: 1 if household belongs to a schedule caste or schedule tribe, 0 otherwise.
OBC	Dummy: 1 if household belongs to an 'other backward caste', 0 otherwise.
MOTIV	Dummy: 1 if main respondent answered 'yes' to the question 'is it important for a girl to be educated?'; 0 otherwise.
Village characteristics	
VVEC	Dummy: 1 if village has a Village Education Committee, 0 otherwise.
VDEVELOP	Village development index: the dummy variables VELEC (village has electricity), VPOST (village has a post office), VPIPEW (village has piped water), and VPHONE (village has a phone) each take the value 1 if the village has the relevant facility and 0 otherwise. Then VDEVELOP = VELEC+VPOST+VPIPEW+VPHONE.
VWASSOC	Dummy: 1 if village has a *mahila mandal* (women's association), 0 otherwise.
School characteristics	
LUNCH	Dummy: 1 if the school provides a mid-day meal, 0 otherwise.
CLASSROOMS	Number of primary classrooms per 100 children aged 6-11 in the village (an indicator of school infrastructure).
WATERPROOF	Dummy: 1 if the school building is water-proof, 0 otherwise.
DAYS6	Dummy: 1 if school was open for 6 days out of the past 7 days, 0 otherwise.
NTDUTIES	Number of days spent by the class-1 teacher in non-teaching duties in the preceding 4 weeks.
PTCOOP	Index of parent-teacher cooperation, constructed as an unweighted sum of four dummy variables: PATTIT (headteacher considers that the attitude of parents towards the school is 'helpful'); PSUPPORT (headteacher approached the parents for help in the preceding year); PCOOP (parents' response to request for help was positive); and PTA (school has a parent-teacher association).
CTRATIO	'Child-teacher ratio': number of children aged 6-11 divided by number of teachers appointed in primary sections. The number of children aged 6-11 was calculated as 0.14*VPOP, where VPOP is the village population and 0.14 is the 1991 census estimate of the proportion of the population in this age group in the PROBE states.
CTRATIO (predicted)	Predicted value of child-teacher ratio, using 'distance to nearest road' and 'village size' as identifying instruments.

Table 3
Binary logit of initial enrolment (5-18 years)

Variable	All			Boys			Girls		
	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect
_cons	-5.783	-7.50 **	-0.488	-6.452	-5.88 **	-0.372	-4.273	-4.74 **	-0.503
age_ch	1.066	9.44 **	0.090	1.506	9.01 **	0.087	0.749	5.62 **	0.088
agesq	-0.047	-9.66 **	-0.004	-0.065	-9.11 **	-0.004	-0.033	-5.40 **	-0.004
edu_mo	0.136	2.98 **	0.012	0.025	0.38	0.001	0.185	2.61 **	0.022
edu_fa	0.112	5.26 **	0.009	0.079	2.31 **	0.005	0.144	6.54 **	0.017
caslab	-0.123	-0.65	-0.010	-0.296	-1.09	-0.017	0.015	0.06	0.002
job	-0.009	-0.03	-0.001	-0.035	-0.10	-0.002	0.045	0.11	0.005
asset	0.032	3.10 **	0.003	0.040	1.59	0.002	0.032	3.13 **	0.004
pccrms	0.031	0.70	0.003	0.054	0.83	0.003	-0.001	-0.01	0.000
cowgoat	-0.012	-0.58	-0.001	0.019	0.74	0.001	-0.032	-1.52	-0.004
landown	-0.010	-0.61	-0.001	0.019	0.61	0.001	-0.028	-1.28	-0.003
depend	-0.213	-2.63 **	-0.018	0.134	0.93	0.008	-0.439	-3.92 **	-0.052
muslim	-0.158	-0.62	-0.013	-0.215	-0.55	-0.012	-0.339	-0.97	-0.040
scst	-0.424	-2.27 **	-0.036	-0.252	-1.03	-0.015	-0.741	-2.95 **	-0.087
obc	-0.426	-2.09 **	-0.036	-0.323	-1.37	-0.019	-0.657	-2.47 **	-0.077
motiv	1.280	5.97 **	0.108	1.161	4.16 **	0.067	1.524	5.73 **	0.179
vvec	-0.014	-0.08	-0.001	-0.046	-0.20	-0.003	0.059	0.23	0.007
vdevelop	0.198	2.70 **	0.017	0.059	0.70	0.003	0.258	2.69 **	0.030
vwassoc	0.347	1.11	0.029	0.097	0.31	0.006	0.576	1.37	0.068
lunch	-0.096	-0.29	-0.008	-1.124	-2.90 **	-0.065	0.972	2.65 **	0.114
classrooms	0.071	1.41	0.006	-0.014	-0.18	-0.001	0.156	2.58 **	0.018
waterproof	0.535	2.57 **	0.045	0.158	0.60	0.009	0.889	3.62 **	0.105
days6	0.331	2.07 **	0.028	0.174	0.82	0.010	0.393	1.91 *	0.046
ntduties	-0.028	-1.25	-0.002	-0.044	-2.44 **	-0.003	-0.023	-0.71	-0.003
ptcoop	0.099	1.39	0.008	0.111	1.21	0.006	0.102	1.00	0.012
ctratio (predicted)	-0.006	-3.92 **	-0.001	-0.009	-2.94 **	0.000	-0.006	-2.41 **	-0.001
male	1.361	8.37 **	0.115						
Log L		-1112.93			-512.02			-542.94	
Restricted Log L		-1514.38			-671.15			-799.66	
Pseudo R-square		0.2651			0.2371			0.3210	
N		3230			1806			1424	
Mean of dep variable		0.8217			0.8776			0.7507	

* significant at 10% level ** significant at 5% level

Table 4: Binary logit of current enrolment (5-12 years)

Variable	All			Boys			Girls		
	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect	coefficient	robust t-value	marginal effect
_cons	-0.279	-0.56	-0.031	1.107	1.57	0.094	-0.692	-1.20	-0.103
age5	-2.471	-8.73 **	-0.271	-2.758	-6.84 **	-0.233	-2.126	-5.25 **	-0.317
age7	0.129	0.51	0.014	0.292	0.91	0.025	0.092	0.22	0.014
age8	-0.204	-1.03	-0.022	0.476	1.52	0.040	-0.678	-2.12 **	-0.101
age9	-0.233	-0.85	-0.025	-0.168	-0.43	-0.014	-0.197	-0.51	-0.029
age10	-0.464	-2.04 **	-0.051	-0.221	-0.71	-0.019	-0.788	-2.38 **	-0.117
age11	-0.610	-2.39 **	-0.067	-0.345	-0.81	-0.029	-0.881	-2.41 **	-0.131
age12	-0.777	-3.18 **	-0.085	-0.549	-1.54	-0.046	-1.178	-2.71 **	-0.175
edu_mo	0.128	2.97 **	0.014	0.052	1.00	0.004	0.170	2.41 **	0.025
edu_fa	0.114	5.60 **	0.013	0.107	3.78 **	0.009	0.132	5.10 **	0.020
caslab	-0.243	-1.11	-0.027	-0.206	-0.74	-0.017	-0.314	-1.13	-0.047
job	0.045	0.16	0.005	0.207	0.50	0.017	-0.001	0.00	0.000
asset	0.034	3.27 **	0.004	0.024	1.08	0.002	0.038	3.62 **	0.006
pccrms	0.033	0.64	0.004	0.047	0.68	0.004	0.017	0.23	0.003
cowgoat	-0.044	-1.55	-0.005	-0.011	-0.27	-0.001	-0.076	-2.04 **	-0.011
landown	-0.001	-0.04	0.000	0.011	0.40	0.001	-0.003	-0.14	0.000
depend	-0.208	-1.95 *	-0.023	-0.087	-0.52	-0.007	-0.333	-2.22 **	-0.050
muslim	-0.107	-0.38	-0.012	-0.282	-0.71	-0.024	-0.103	-0.28	-0.015
scst	-0.594	-2.73 **	-0.065	-0.515	-1.74 *	-0.044	-0.802	-2.94 **	-0.119
obc	-0.577	-2.70 **	-0.063	-0.500	-1.82 *	-0.042	-0.780	-2.98 **	-0.116
motiv	1.366	6.03 **	0.150	1.145	3.66 **	0.097	2.033	7.24 **	0.303
vvec	0.132	0.63	0.014	0.155	0.60	0.013	0.075	0.26	0.011
vdevelop	0.182	2.36 **	0.020	0.069	0.72	0.006	0.244	2.37 **	0.036
vwassoc	0.306	1.17	0.033	0.040	0.13	0.003	0.687	1.97 **	0.102
lunch	0.092	0.31	0.010	-0.626	-1.61	-0.053	0.998	2.89 **	0.149
classrooms	0.028	0.50	0.003	0.000	0.01	0.000	0.054	0.74	0.008
waterproof	0.586	2.81 **	0.064	0.321	1.23	0.027	0.979	3.33 **	0.146
days6	0.384	2.19 **	0.042	0.015	0.07	0.001	0.747	3.17 **	0.111
ntduties	-0.011	-0.58	-0.001	-0.006	-0.30	0.000	-0.016	-0.54	-0.002
ptcoop	0.047	0.69	0.005	0.044	0.41	0.004	0.056	0.58	0.008
ctratio (predicted)	-0.004	-1.80 *	0.000	-0.003	-0.92	0.000	-0.005	-2.13 **	-0.001
male	1.245	7.80 **	0.136						
Log L		-915.12			-461.60			-412.76	
Restricted Log L		-1289.95			-613.07			-644.78	
Pseudo R-square		0.2906			0.2471			0.3598	
N		2472			1405			1067	
Mean of dep variable		0.7840			0.8420			0.7076	

Table 5
Ordered logit model of grade attainment (13-18 year olds)
(no schooling = 0; 1-4 years of schooling = 1; >=5 years of schooling = 2)

Variable	All			Boys			Girls		
	coefficient	robust t-value	marginal effect p(>=grade5)	coefficient	robust t-value	marginal effect p(>=grade5)	coefficient	robust t-value	marginal effect p(>=grade5)
age14	0.491	1.69 *	0.081	0.398	0.91	0.040	0.533	1.40	0.119
age15	0.085	0.31	0.011	-0.350	-0.88	-0.044	0.309	0.83	0.069
age16	0.234	0.78	0.035	0.293	0.63	0.029	-0.152	-0.40	-0.034
age17	0.102	0.39	0.013	-0.092	-0.28	-0.016	0.370	0.89	0.083
edu_mo	0.364	3.33 **	0.064	0.327	1.47	0.037	0.387	2.98 **	0.087
edu_fa	0.115	3.40 **	0.020	0.117	2.16 **	0.014	0.140	3.72 **	0.031
caslab	-0.051	-0.17	0.000	-0.399	-0.91	-0.038	0.448	1.39	0.100
job	0.021	0.05	0.000	-0.250	-0.42	-0.031	-0.025	-0.04	-0.006
asset	0.012	1.45	0.002	0.000	0.03	0.000	0.021	2.04 **	0.005
pccrms	0.039	0.65	0.006	0.324	2.43 **	0.036	-0.063	-0.96	-0.014
cowgoat	0.070	2.04 **	0.012	0.203	2.94 **	0.022	-0.007	-0.15	-0.002
landown	-0.033	-1.69 *	-0.006	-0.035	-1.02	-0.004	-0.031	-1.03	-0.007
depend	-0.141	-1.28	-0.028	0.185	0.99	0.017	-0.438	-2.81 **	-0.098
muslim	-0.275	-0.86	-0.037	-0.712	-1.68 *	-0.069	-0.171	-0.35	-0.038
scst	-0.485	-1.58	-0.089	-0.382	-0.85	-0.046	-0.538	-1.47	-0.120
obc	0.177	0.74	0.027	0.733	2.00 **	0.078	-0.296	-0.97	-0.066
motiv	0.876	3.07 **	0.159	1.019	3.00 **	0.119	0.592	1.44	0.132
vvec	-0.169	-0.69	-0.031	-0.218	-0.70	-0.026	-0.056	-0.16	-0.013
vdevelop	0.202	1.81 *	0.036	0.233	1.37	0.026	0.262	1.84 *	0.059
vwassoc	1.088	2.65 **	0.192	0.762	1.35	0.086	1.200	2.87 **	0.269
lunch	0.444	1.35	0.076	-0.033	-0.06	-0.006	1.336	2.73 **	0.299
classrooms	0.050	0.77	0.010	-0.037	-0.34	-0.002	0.107	1.51	0.024
waterproof	0.599	2.01 **	0.103	-0.017	-0.04	-0.005	1.164	3.42 **	0.260
days6	-0.054	-0.22	-0.005	0.029	0.09	0.008	-0.074	-0.23	-0.017
ntduties	-0.035	-1.13	-0.006	-0.069	-2.33 **	-0.008	-0.007	-0.21	-0.002
ptcoop	0.390	3.91 **	0.068	0.516	3.77 **	0.058	0.333	2.32 **	0.074
ctratio (predicted)	-0.008	-3.48 **	-0.002	-0.010	-3.52 **	-0.001	-0.008	-2.35 **	-0.002
male	1.580	7.47 **	0.281						
Log L		-509.92			-226.47			-255.04	
Restricted Log L		-658.11			-293.54			-349.58	
Pseudo R-square		0.2252			0.2285			0.2704	
N		760			402			358	

Table 6
Mean values of variables: Himachal Pradesh vs full sample

Variable	HP mean	Sample mean	marginal effect	'Shift effect'[a]
age5	0.107	0.142	-0.317	+0.011
age7	0.145	0.163	0.014	-0.000
age8	0.122	0.159	-0.101	+0.004
age9	0.122	0.105	-0.029	- 0.000
age10	0.122	0.103	-0.117	+0.002
age11	0.084	0.084	-0.131	- 0.000
age12	0.153	0.111	-0.175	- 0.007
edu_mo	3.649	1.522	0.025	+0.053
edu_fa	7.298	5.590	0.020	+0.034
caslab	0.069	0.151	-0.047	+0.004
job	0.221	0.114	0.000	+0.000
asset	11.786	10.336	0.006	+0.009
pccrms	0.951	1.417	0.003	- 0.001
cowgoat	2.652	3.437	-0.011	+0.009
landown	2.422	3.789	0.000	- 0.000
depend	1.329	1.506	-0.050	+0.009
muslim	0.015	0.093	-0.015	+0.001
scst	0.496	0.314	-0.119	- 0.022
obc	0.153	0.306	-0.116	+0.018
motiv	0.962	0.918	0.303	+0.013
vvec	0.420	0.570	0.011	- 0.002
vdevelop	3.328	1.483	0.036	+0.066
vwassoc	0.626	0.156	0.102	+0.048
lunch	0.000	0.086	0.149	- 0.013
classrooms	2.357	1.239	0.008	+0.009
waterproof	0.630	0.351	0.146	+0.041
days6[b]	0.049	0.354	0.111	- 0.034
ntduties	0.695	1.916	-0.002	+0.002
ptcoop	2.109	1.398	0.008	+0.006
ctratio[c]	43.014	68.904	-0.001	+0.026
N	131	1067		

[a] Difference between the HP mean and the sample mean, multiplied by the estimated 'marginal effect' on current enrolment for girls (reproduced in the preceding column, from Table 4).

[b] The low value for HP is misleading: there were several local holidays during the survey period.

[c] The shift effect for this variable is calculated by combining the marginal effect associated with 'predicted CTRATIO' (see Table 4) with the relevant means of the actual CTRATIO, reported in the second and third column; the corresponding means of the predicted CTRATIO are 24.0 (in HP) and 69.7 (for the whole sample), respectively.

Note: The means in this table are calculated using all girls aged 5-12 as the reference group.

[1] Calculated from International Institute for Population Sciences (1995), p.56. The reference year is 1992-93.

[2] The acronym PROBE refers to the Public Report on Basic Education where the main findings of the survey were presented (The PROBE Team, 1999). Both of us were personally involved in the survey, in collaboration with other researchers. In that sense, this study may be considered as an exercise in 'participatory econometrics' (Rao, 1998).

[3] The last study is particularly relevant; a brief comparison of Sipahimalani's (1998) results with our own is presented in

an earlier, more detailed version of this paper (Drèze and Kingdon, 1999).

[4] These four states were selected partly for convenience (they share a common language, Hindi), and partly because they are India's most educationally backward states. Together, they account for 40% of India's population, and over half of all out-of-school children. Aside from these four major states, the PROBE survey also covered the smaller state of Himachal Pradesh (see below).

[5] This sketch is based on the findings reported in The PROBE Team (1999); it is consistent with other field studies as well as with secondary data. For a useful survey of field-based investigations of schooling in India, see Bhatty (1998).

[6] These figures are higher than those quoted in the opening paragraph, suggesting rapid improvement in school participation between 1992-3 and late 1996 (when the PROBE survey took place).

[7] The foundations of this success are explored in The PROBE Team (1999), chapter 9. The regression results presented in this paper throw some further light on the issue, as discussed in section 7.

[8] A child is taken to 'participate' in the schooling system if he or she is reported by her parents to be enrolled in a school. The terms 'school participation' and 'school attendance' will be used interchangeably.

[9] Scheduled castes are at the bottom of the traditional caste hierarchy and are entitled to various forms of positive discrimination. In government schools, for instance, scheduled-caste children are sometimes eligible for free textbooks or other incentives.

[10] Kingdon's (1998) estimation procedure corrected for possible 'selection bias'.

[11] For earlier analyses of the relation between teacher-pupil ratios and school participation (or pupil achievements) in India, see Heyneman and Loxley (1982, 1983) and Kingdon (1994, 1996).

[12] In an analysis of survey data for Lucknow city, Kingdon (1996) finds no relation between teacher salaries and pupil achievements after controlling for teacher education and training as well as for pupil, parental and school characteristics (on this, see also Fuller, 1986). Teacher salaries in India depend mainly on years of service.

[13] Current enrolment refers to the status of the child during the relevant year (i.e. in 1996-7), not to school attendance on the specific day of the survey.

[14] In the states concerned, the primary stage consists of grades 1-5. This roughly corresponds to the 6-10 age group, though some children are enrolled in grade 1 at age 5, and many are still in primary school at the age of 11 or even 12.

[15] The number of children in each village was estimated as a constant ratio of the total village population. The constant is the share of the relevant age group in the total population of the PROBE states, according to 1991 census data.

[16] Another consideration reduces the endogeneity problem: the 40:1 norm, when it is applied at all, is based on school enrolment data, which are known to be fudged. The different ways in which fudging takes place are discussed in the PROBE report (The PROBE Team, 1999, pp.91-92). For present purposes, it is sufficient to note that, in many cases, the number of children listed in the school registers is probably closer to the total number of children in the village than to the actual number of children attending school. This, again, would imply that teacher postings and child-teacher ratios are not very sensitive to school participation rates.

[17] Both variables are exogenous, and there is no reason why they should have much of a direct effect on school participation after controlling for school facilities, village development and related variables. Child-teacher ratios, on the other hand, are positively related to 'distance from the nearest road' (because teachers dislike remote postings), and also to 'village size' (because extra teacher postings appear to under-compensate for increases in population size). Exogeneity tests are somewhat inconclusive: a Smith-Blundell test of the exogeneity of CTRATIO rejects the null hypothesis at the 10% level.

[18] The diversion of 'non-teaching duties' is a common complaint of teachers in rural India. Teachers are often mobilised, for instance, to help with the decennial census, the cattle census, vote counting, health programmes and literacy campaigns (see The PROBE Team 1999, chapter 5).

[19] Two variables called for a somewhat different aggregation rule: (1) in the case of CTRATIO, we simply divided the estimated child population of the village by the total number of primary-grade teachers in the village; (2) similarly, the CLASSROOMS variable simply measures the total number of primary classrooms per 100 children in the village.

[20] The corresponding question for boys ('Is it important for a boy to be educated?') is of little use here, since 99% of the respondents answered 'yes'.

[21] In these tables, we report robust t-values adjusted for cluster effects, i.e. the possibility of correlated errors across individual observations within each village (Deaton, 1997, p.77; Moulton, 1990).

[22] The influence of male observations in the pooled regressions is somewhat greater in the 'grade attainment' regressions (Table 5). This is not surprising, since the variation in grade attainment among boys is much greater than the variation in 'initial enrolment' or even 'current enrolment' (also among boys).

[23] One particular version of the latter view is that the gender gap in educational achievements reflects a rigid equilibrium condition of the 'marriage market'.

[24] The marginal effects of the age dummies in Table 4 yield interesting information on drop-out rates; see Drèze and Kingdon (1999) for further discussion.

[25] These patterns (and most other findings) also hold if MOTIV is dropped from the regression.

[26] As mentioned earlier, this also applies to various other indicators of transient aspects of school quality, not included in Tables 3 to 5.

[27] Glewwe and Jacoby (1994) report a similar finding for Ghana, where leaky classrooms are associated with low cognitive achievements. The authors even argue that their estimates (combined with back-of-the-envelope cost calculations) have "uncovered the relative effectiveness of repairing school buildings over investments in instructional materials, such as books, desks and blackboards and in teacher quality" (pp. 862-3).

[28] Other kinds of community effects are possible. For instance, bearing in mind that VDEVELOP is largely an index of local public services (see Table 2b), it can also be interpreted as an indication of the political clout of the relevant village, or of the quality of local governance, or of cohesion in the village community - all of which may lead to improvements in school quality not otherwise captured in the data.

[29] Note that VECs are recent institutions. This variable, therefore, is somewhat irrelevant in the case of 'initial enrolment' as well as 'grade attainment'.

[30] In this connection, it is worth noting that female schooling is far more responsive than male schooling to the economic status of the household (as captured here by ASSET). For a similar finding in Pakistan, see Jensen (1999).

[31] A more sophisticated decomposition of Himachal Pradesh's advantage, allowing for different regression coefficients in HP and elsewhere (e.g. using Oaxaca's method), is difficult to carry out due to the small number of observations for Himachal Pradesh. Note also that the shift effects reported in Table 6 are based on linear approximation: strictly speaking, the logit model is inherently non-linear.