Forthcoming in the ***Economics of Education Review***

Education and poverty in rural China

Phil Brown and Albert Park
Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
U.S.A.

May 2001

Abstract

We analyze household and school survey data from poor counties in six Chinese provinces to examine the effects of poverty, intra-household decision-making, and school quality on educational investments (enrollment decisions) and learning outcomes (test scores and grade promotion). Unlike previous studies, we use direct measures of credit limits and women's empowerment. Drawing a distinction between the effects of wealth (measured by expenditures per capita) and credit constraints, we find that the former improves learning while the latter reduces educational investments. We find evidence of a story of gender bias in which academically weak girls are more likely to drop out in primary school while most boys continue on to junior secondary school. Women's empowerment reduces the likelihood of dropping out but does not affect other outcomes. Finally, our measures of school quality have some effect on the duration of primary school enrollment but not on learning.

JEL Classifications: D1, I2, O1
Keywords: demand for schooling, economic development, human capital

Received 5 November 2000; accepted 15 March 2001

+Corresponding author: Albert Park, alpark@umich.edu, T: 734-764-2363, F: 734-764-2769.

1. Introduction

In developing countries, poverty is often associated with low levels of educational attainment, as well as larger gender gaps in education (Filmer 2000). Low incomes and wealth combined with incomplete credit markets make it difficult to finance educational investments even when the returns exceed the costs. In addition, even after controlling for wealth differences, a robust finding is that parents with lower levels of education are less likely to educate their own children. Poorly educated parents may value education less, may have low scholastic aptitude which they pass on to their children, or may be less able to provide complementary inputs to learning (e.g., helping children with homework). The low education of mothers, in particular, may reduce their bargaining power within the household and affect family educational decisions if parental preferences over education differ. Further, a lack of community resources in poor areas often leads to lower quality schools, which may reduce the returns to education and discourage enrollment. Finally, in segmented labor markets, the returns to education in poor, remote areas may be sufficiently low to discourage educational investments.

In this paper, we analyze data collected from surveys of households and schools in poor counties in six Chinese provinces to examine the effects of individual, family, and school characteristics on educational attainment, focusing in particular on the importance of poverty and credit constraints, intra-household decision making (especially as it relates to gender), and school quality. The detailed data make possible several innovations. First, unlike many studies that focus on single measures of attainment, especially enrollment, we examine multiple outcomes that reflect both investments in schooling and learning within school. Next, the data enable us to construct more direct measures of variables of interest than in previous studies. Using data on existing debt and the ability to borrow money from both formal and informal sources, we construct a measure of household credit limits, which allows us to test separately the effects of wealth and credit limits. A direct question on the role of mothers relative to fathers in the decision to enroll children in school serves as a measure of women's empowerment. Last, local school quality measures, not included in most household surveys, are available from separate surveys of local primary and junior secondary schools. These innovations provide general insights, but also enable us to go well beyond the scope of existing empirical studies of educational attainment in rural China, which have typically used large data sets with limited information.

Previous research suggests that the different hypothesized connections between poverty and educational attainment are likely to be important in the Chinese context. Tsang (1996) and Hannum (1998) report that many schools have increased fees to offset rising costs resulting from education decentralization, and Park and Wang (2000) find that twelve percent of informal loans to households in our sample are used to pay school fees, which implies that credit constraints may be important for some poor households. Hossain (1996) reports that the poorest quintile of households in China spend 14.2 percent of annual income on education, while the wealthiest quintile spend only 5.5 percent.

With regard to intra-household decision-making, Knight and Song (2000) use 1995 survey data to show that a wife's bargaining position, measured by the relative education level of the mother, is positively correlated with children's education, and disproportionately so for boys. They also find that boys have a higher probability of enrollment at all levels. Hannum (1998) uses census data to demonstrate that boys are more likely to enroll than girls, and that this gap is exacerbated when households face resource constraints. Thus, it is not surprising that the gender gap in enrollment is much larger in poor counties than in non-poor counties (World Bank, 1999). Research also finds that after controlling for wealth and expenditure levels, educated parents in China are more likely to educate their children (Jamison and van der Gaag, 1987; Connelly and Zheng, 2000).

Finally, school quality is likely to be a concern in the Chinese context. China's fiscal system has struggled to generate adequate revenue, leading to a marked decentralization of fiscal responsibility and a revenue crisis for governments in poor counties (Park, Rozelle, Wong, and Ren, 1996). This has led to large differences in public spending on education and in teacher quality across regions (Tsang, 1994; West, 1996). World Bank (1999) reports that the recurrent per-pupil expenditure in the wealthiest 10 percent of counties was more than 4.5 times that in the poorest 10 percent in 1997. Unfortunately, no existing studies of educational attainment in China empirically examine the effects of school quality.

The remainder of the paper is organized as follows. Section 2 presents a simple model to illustrate the interplay of credit constraints, intra-household decision-making, and school quality variables in educational investment decisions. Section 3 describes the 1997 survey and describes the dependent and independent variables used in the analysis. Section 4 introduces China's rural educational system. Section 5 describes the empirical specifications and identification strategy. Section 6 presents descriptive and estimation results for each of the educational attainment outcomes. Section 7 concludes.

2. Modeling educational investments

We model the educational investment decision made by a family consisting of a mother, a father, and a single child. Educational investments (i.e., the number of years of schooling) are made by parents, who maximize a joint utility function U which is a weighted sum of parent and child payoffs.[1] Household income during the period of investment (and any initial wealth) is y, and the family invests E_c in the child's education and must pay a cost of P_E, which includes required school fees as well as the opportunity cost of the child's time (which we assume accrues to parents). Let R denote the returns to the child's education, and let α be the share of the returns that are transferred from the child to the parents through future financial support and care. Thus, $(1-\alpha)$ is the share of returns retained by the child. The parameter A represents the degree to which parents are altruistic toward their children. If $A=1$, parents care as much about their children as themselves. Total spending on education cannot exceed the sum of income and the household's credit limit $(\bar{b})$. The parent's utility maximization problem is thus:

$$\underset{E_c}{Max}\, U = y - P_E E_c + \alpha R(E_c) + A(1-\alpha)R(E_c) \qquad (1)$$
$$s.t.\; P_E E_c \le y + \bar{b}$$

where $\alpha \in [0,1]$, $A \in [0,1]$, and $\bar{b} \ge 0$.[2] To simplify, we assume a zero interest rate and perfect enforcement of lending contracts.

We make several further assumptions about the model parameters. First, the share of returns to education retained by the parents is a function of the child's sex, i.e., $\alpha = \alpha(S)$. This is plausible in that in China daughters marry and leave the family, while sons often co-reside with parents and are generally responsible for the support of elderly parents (Parrish and Willis, 1993; Hannum and Xie, 1994; Hannum, 1998). Second, we model altruism as a linear combination of mother's preferences (A_m) and father's preferences (A_f), the relative weight placed on each depending upon the mother's intra-household bargaining power (β). Parental preferences are a weighted combination of sex (S), and parental education $(E_m \text{ and } E_f)$. Thus, the altruism parameter is defined as follows:

$$\begin{aligned} A &= \beta A_m + (1-\beta) A_f \\ A_m &= a_1 \mathrm{S} + a_2 \mathrm{E}_m \\ A_f &= b_1 \mathrm{S} + b_2 \mathrm{E}_f \end{aligned} \qquad (2)$$

where $\beta \in [0,1]$. Thus, $A = A(\beta, \mathrm{S}, \mathrm{E}_{\mathrm{m}}, \mathrm{E}_f)$.

Assuming a Cobb-Douglas function with decreasing returns, the return function can be expressed as:

$$R(E_c) = rE_c^{\phi} \qquad (3)$$

where $0 < \phi < 1$. Let **X** be a vector of variables that affect the returns to schooling, so that $r = \mathbf{w}'\mathbf{X}$, where **w** is a coefficient vector. The variables affecting returns to education, **X**, include child characteristics ($\mathbf{X_C}$), household characteristics ($\mathbf{X_H}$), and school characteristics ($\mathbf{X_Q}$). Thus, $\mathbf{X} = \lfloor \mathbf{X_C}\ \mathbf{X_H}\ \mathbf{X_Q} \rfloor$.

If credit constraints do not bind, then the first order condition for (1) is:

$$\frac{\partial R}{\partial E_c} = \frac{P_E}{\alpha + (1-\alpha)A} \qquad (4)$$

For the Cobb-Douglas return function, we can solve explicitly for the unconstrained optimum:

$$E_c^U = \left[\frac{[\alpha + (1-\alpha)A]\phi \mathbf{w}'\mathbf{X}}{P_E} \right]^{\frac{1}{1-\phi}} \qquad (5)$$

Note that if parents capture the entire return to children's education (α=1) or if parents are fully altruistic (A=1), then the first order condition (4) collapses to $R'(E_c) = P_E$, i.e., the marginal return equals the price. This special case serves as an efficiency benchmark (denoted E_c^E) since the investment decision maximizes social returns. Educational investments are only affected by factors that affect returns.

If the credit constraint does bind, however, then the constrained optimum is

$$E_c^C = \frac{y + \bar{b}}{P_E} \qquad (6)$$

In this case, educational investments are solely determined by income and credit limits.

Thus, under different assumptions (unconstrained, efficient, and constrained) educational investments are functions of different arguments:

$$\begin{aligned} E_c^U &= E_c^U[\mathbf{X}, P_E, A(S, E_f, E_m, \beta), \alpha(S)] \\ E_c^E &= E_c^E[\mathbf{X}, P_E] \\ E_c^C &= E_c^C[y, \bar{b}, P_E] \end{aligned} \qquad (7)$$

Note that **X** contains the full set of independent variables and that all variables affect whether the credit constraint binds, so that these different functions do not provide overriding restriction tests to distinguish among E_c^U, E_c^E, and E_c^C. But they do illustrate the multiple pathways through which variables of interest may affect educational outcomes, and so facilitate interpretation of the estimation results.

Consider, for example, the effect of a child's sex, which may be important if the returns to education differ for boys and girls, either because of labor market conditions, differential treatment in school, different levels of motivation, or different support for educational attainment at home. In addition, the share of the returns to education accruing to parents may differ by sex if girls marry and leave the family while boys remain within the family after marriage. Finally, the altruism that parents show to their children may differ for sons and daughters. If the preferences of fathers and mothers differ, bargaining power within the household also matters.

Household characteristics also affect schooling. Economic variables, including y and $\bar{b}$, impact educational investments by facilitating the purchase of goods that are complementary to learning (e.g., food, utilities, furniture) and, when credit constraints bind, directly determining the ability of households to finance desirable educational investments. Parental education affects optimal schooling levels by increasing returns (e.g., if educated parents provide more or better support for children's learning or have connections to better jobs in the labor market) and by affecting altruistic preferences (which depend on the interaction with women's empowerment if parental preferences differ). Education of parents also increases educational investments in children indirectly through household income and expenditures, and women's empowerment. Finally, school quality affects educational attainment by increasing the returns to education.

3. Data and variables

The data come from a 1997 survey of households conducted by one of the authors in collaboration with the China Poverty Research Association. The households are located in six poor counties, each in a different province: Shaanxi and Gansu in the northwest, Sichuan and Guizhou in the southwest, and Henan and Jiangxi in central China. The provinces were chosen to broadly represent different poverty regions in the country. The county chosen in each province was selected from among counties that were: 1) nationally designated poor counties; 2) State Statistical Bureau (SSB) national rural household sample survey counties (about one third of all counties in China); and 3) located in the main poverty belt within each province. The household sample in each county was the same as that selected by the SSB, which draws a nationally representative stratified random sample each year. The survey encompassed 446 households and the 40 primary schools and 37 junior secondary schools that serve them. School data come from interviews with local primary and junior secondary school principals, which included questions about school infrastructure, teachers, enrollment, and finances. Student test scores in the most recent semester were also collected.

The household part of the survey included 472 school-aged children (between five years, six months and 16 years, 11 months). Of these, 296 were enrolled in primary school, 71 were in junior secondary school, and 3 were in senior secondary school (Figure 1). There were 55 drop outs, and 47 children had never enrolled. Of those who never enrolled, 83.0 percent were below age 10 at the time of the survey and thus plausibly would enroll in the future. Households provided data on time allocation, assets, income and credit, and family background. Table 1 presents summary statistics for the households with children and schools in the sample. Mean per capita expenditure is 1134 yuan (in 1997, US$1 ≅ 8 yuan), the mean household credit limit is 4643.4 yuan, the mean number of children is 2.2, and the mean number of years of parental education is 7.3 years for fathers and 3.4 years for mothers. Data on school-related variables are described below.

[Figure 1 HERE]

[Table 1 HERE]

3.1 Dependent variables

We study the determinants of one educational investment measure (years of schooling) and two learning outcomes (test scores and whether the child was ever held back). The former correspond to E_c^C in the model, the latter to the return function, R. Differences in labor market returns are controlled for by imposing community fixed effects.

Years of schooling are calculated as the sum of grades completed and years held back. Examination scores are the average scores on the most recent language and math exams, which are administered each semester and which are the same for students in the same grade in the same county. We standardize test scores by grade within each county to make grades on different tests comparable; test score is thus defined as the number of standard deviations from the mean score of all children in the same grade in the same county. The survey also asks about the number of years held back, but does not report the grades in which children were held back. The large majority of those ever held back are held back for one year only (74 percent). We thus focus attention on whether children have ever been held back.

3.2 Independent variables

As described above, factors affecting educational investments and learning outcomes (**X**) include child ($\mathbf{X_C}$), household ($\mathbf{X_H}$), and school quality variables ($\mathbf{X_Q}$). $\mathbf{X_C}$ includes sex (S), age of enrollment, number of older siblings, and number of younger siblings.[3] $\mathbf{X_H}$ includes household expenditures per capita (y), the household's credit limit $(\bar{b})$, father's education (E_f), mother's education (E_m), and women's empowerment (β), which is also interacted with mothers education and child gender. $\mathbf{X_Q}$ includes the student-teacher ratio, the percentage of classrooms that are rainproof, the percentage of teachers with post-secondary education, and under certain assumptions, school fees and distance to school. School fees and distance to school are also measures of the price of educational investments (P_E).

Expenditures per capita, calculated from self-recorded diaries kept by households and tabulated by local State Statistical Bureau enumerators, is our main poverty measure. With incomplete capital markets, expenditures per capita reflect wealth effects (Glewwe and Jacoby, 2000), which often have been found to influence educational investment decisions.[4] In theory, wealth could affect educational decisions even when credit constraints are not binding if wealthier households consume goods that are complementary to learning and also provide consumption value to the household for reasons unrelated to education (e.g., nutritious food, tables and chairs, books, TV). Thus, non-separable consumption and educational investment decisions can lead to wealth effects even in the absence of credit constraints. Tests using direct measures of credit constraints can help clarify the ambiguity inherent in measured wealth effects. In the survey, respondents were asked the value of outstanding formal and informal loans, and the additional amount that they felt they could borrow either from institutions or from friends and family members in the event of an emergency. Our credit limit variable is the sum of these values. Theory says that wealth should have a strong effect only when credit constraints bind; therefore we generate an interaction dummy variable indicating whether households are both poor and credit constrained, defined as being below the 33rd percentile of the sample in terms of both expenditure per capita and credit limits, accounting for 14.1 percent of sampled households.[5]

In evaluating the decision-making role of women versus men, the survey asks which parent is responsible for deciding whether children attend school. The variable for women's empowerment equals one if the wife decides, 0.5 if both decide, and zero if the husband decides. The definition of altruism in (2) suggests that women's empowerment should be interacted with the child's gender and mother's education.[6]

Conditional on ability, previous achievement, and earnings prospects, school quality has been found to have a positive impact on enrollment in other studies (e.g., Hanushek and Lavy, 1994). We focus on variables that measure different key aspects of school quality – class size, teacher quality, and infrastructure. Following much of the literature, our specific measures are the student-teacher ratio, the percentage of teachers with post-secondary education,[7] and the percentage of classrooms that are rainproof (Glewwe and Jacoby, 1994). In addition to these measures, our two cost of schooling measures, school fees and distance to school, might also reflect differences in school quality. Villages that set higher school fees may have larger per-pupil budgets, and field interviews suggest that schools that serve multiple villages (i.e., schools that are likely to be farther away) achieve economies of scale and receive better funding.

4. China's rural educational system

While the minimum age of enrollment in China is six, households in many areas are accustomed to sending their children to school at older ages. The mean age of enrollment in our sample is 7.4, or about one year later than would be expected if all children enrolled as soon as possible after age six.[8] As seen in Figure 2A, a significant proportion of children do not start school until they are 8 or older, and girls are more likely to start later (the average starting ages are 7.3 for boys and 7.5 for girls). Interestingly, the age of enrollment for junior secondary school is lower for girls (mean of 13.1 versus 13.5 for boys, see Figure 2B).[9] This finding suggests a selection process in which only academically strong girls stay enrolled through primary school.

[Figures 2A and 2B HERE]

Nearly all children walk to the nearest primary school, usually located in the village. Primary school is completed in five or six years, depending on the region. Junior secondary schools are usually located in the nearby township. Despite the compulsory education law mandating nine years of education, children whose families do not pay school fees are not allowed to attend school. In our sample, school fees averaged 100.9 yuan in primary schools and 317.8 yuan in junior secondary schools (Table 1). Additional, non-required school-related fees, e.g., supplies and books, averaged 71.4 yuan per child. Thus, a family with one child in primary school and another in junior secondary school would spend about 550 yuan on school-related expenses, or fifty percent of mean expenditures per capita, likely a very high share of a family's cash income.

School quality has emerged as an important concern in China, where fiscal reforms have reduced redistributive budgetary transfers, exacerbating inequities. In our sample of schools, the mean student-teacher ratio is 28.5 for primary schools and 15.1 for junior secondary schools. The mean percentage of teachers with post-secondary education is 54 percent at the primary level, and 88 percent at the junior secondary school level. Seventy-eight percent of primary school classrooms and nearly all junior secondary school classrooms are rainproof. There is significant variation in school quality among provinces (Table 1). For example, just 28 percent of primary school teachers have post-secondary education in Guizhou, compared to 86 percent in Shaanxi.

5. Empirical specification

We analyze the determinants of years of schooling, test scores, and whether children have ever been held back. With exceptions noted below, we include a consistent set of child, household, and school quality variables, as described above.

We model the duration of schooling as a Cox proportional-hazard model (see, for example, Khandker (1996) and Glewwe and Jacoby (2000)). Hazard models account for the dependence of current enrollment on past enrollment decisions, and handle censored observations (students currently enrolled at the time of the survey) in a natural way. The Cox model is attractive because it does not require a parametric specification of the baseline hazard function and thus allows the baseline hazard rate for each community to vary (Cox and Oakes, 1984).[10] We estimate separate hazard models for dropping out of primary school conditional on primary school enrollment, and for dropping out of junior secondary school conditional on junior secondary school enrollment. The hazard ratios can be interpreted as risk multipliers.[11]

Because nearly all children attend at least one year of primary school, there is no selection bias in the sample used to study the primary school duration of schooling. We include age of enrollment as an independent variable in the dropout hazards and other outcome equations because we expect age to affect school performance and the opportunity cost of children's time. We recognize that the coefficient will be upward biased if unobserved poor ability or lack of parental support delays the age of enrollment or makes dropping out more likely.

Test scores are regressed on the full set of independent variables using OLS, with different specifications employing county, village, and household fixed effects. We specify the equation estimating whether a child was ever held back as a conditional logit in order to be able to include county, village, and household fixed effects without introducing bias. We also include dummy variables for the number of grades completed to account for the fact that students who have reached higher grades have more chances to be held back.

Identification

Some independent variables may be endogenous because of simultaneity or omitted variables. Variables resulting from household decisions made by parents, such as expenditures per capita and number of siblings, are particularly susceptible to such bias because they are likely to be made simultaneously with investments in children's education. Expenditures in particular may include educational costs, which naturally increase if children are enrolled. To deal with this specific problem, our expenditure measure excludes educational expenditures, which creates downward rather than upward bias on the expenditure coefficient – a more severe test for finding significant effects. Expenditure levels also reflect household income, which is affected by labor supply decisions of parents, which in turn may depend on whether or not children are in school. Fertility may be negatively correlated with educational investments if there is a tradeoff between quantity and quality of children. However, given China's strict family planning policy, the number of children in many rural families is below desired levels, especially in poor areas.[12] When we regress the number of children on parental education and other parental characteristics, we find no significant effects.

Coefficients on household decision variables and on variables that are plausibly exogenous to household decisions on education (e.g., father's and mother's education, women's empowerment, and credit limits) also may misleadingly pick up the effects of unobserved child and/or parent characteristics. Parental education, for instance, may correlate positively with higher motivation or ability, which may also correlate with willingness to invest in children's education. If this is true, the inclusion of other variables that reflect ability and motivation, such as women's empowerment, expenditures per capita, and credit limits, could better isolate the effect of preferences related to parental education.

Without better data, dealing with all of these endogeneity concerns is challenging. As a practical matter, the vast majority of studies, especially those using cross-sectional data, do not attempt to do so, ignoring potential bias (e.g., Jamison and Lockheed, 1987; Parish and Willis, 1993; Glewwe, Grosh, Jacoby, and Lockheed, 1995; Khandker, 1996; Case and Deaton, 1999; King, Orazem, and Paterno, 1999). A few studies treat income and expenditure data as endogenous (Glewwe and Jacoby, 1994; Glewwe and Jacoby, 2000), and Lillard and Willis (1994) explicitly model the endogeneity of parental education. A conservative approach is to restrict the variable set to those that are strictly exogenous and do not reflect household decisions, leaving out variables such as expenditures per capita. Although this solves the simultaneity problem, it does not solve the omitted variables problem, and the strict reduced form estimates may be difficult to interpret because they are picking up multiple effects. Another approach is to use instrumental variables, but it may be difficult to find suitable instruments that are plausibly exogenous and explain sufficient variation in the endogenous variable. Finally, one can add additional controls to try to pick up background factors, but the possibility of omitted variable bias remains.

In our estimation, we tried a combination of these approaches. Adopting linear specifications for each outcome, we instrumented expenditures per capita and credit limits using cultivated land and the share of cultivated land that is irrigated. To help control for unobserved parental attributes, we added background variables such as the education of grandparents and the number of siblings of each parent. In the end, however, we report estimates from specifications that do not control for endogeneity because none of our alternative specifications substantially alters the magnitude or sign of our coefficient estimates. Our instrumental variables, although significant in first stage regressions, suffer from being "weak" in that they do not explain sufficient variation in the endogenous variable to produce precise estimates (Bound, Jaeger, and Baker, 1995). Nonetheless, inclusion of IVs increased rather than decreased the magnitude of the coefficient of the instrumented variable in every specification, suggesting that our coefficients underestimate the true effects. Including family background variables did not appreciably alter the statistical significance or magnitudes of our estimates, and we dropped them to maximize sample size, since data on background variables were missing for some households. Even if endogeneity bias remains, our estimates are still informative in describing the statistical association between outcomes and various individual, household, and community factors, providing suggestive evidence, if not definitive proof, of causal relationships.

Estimates of the determinants of test scores may be subject to sample selection bias because data are available only for enrolled children. Despite the difficulty of finding convincing identifying instruments, we estimate Heckman selection-correction models of test scores.[13] We find that the selection correction term does not enter significantly into the test score regression but that the effects of gender become smaller in magnitude and statistically insignificant. This lends weak support to the notion that a selection story underlies gender differences in test scores. However, because of the questionable identification assumptions we have imposed, we do not want to read too much into these results. Rather, we present the results for the uncorrected estimates and consider possible biases introduced by selection effects.

Another possible selection problem is endogenous school choice. If children who have higher ability or more supportive parents choose to attend higher quality schools, the measured effect of school quality variables will be biased upward. However, 94.1 percent of the children in our sample attended the nearest primary school, and of those that do not, 59 percent report that the main reason for not doing so is unrelated to school quality. This suggests that only 2 percent of children are changing schools for reasons related to quality.

Even without endogenous school choice, the student-teacher ratio may suffer from endogeneity because it is affected by the enrollment decisions of households, resulting in downward bias. This may be more important in middle school where dropout rates are higher. To deal with this potential problem, we instrument the student-teacher ratio in the test score regression with village population, and find that the results do not change.

In all specifications, we control for community unobservables by including a set of community dummy variables (or, in the hazard estimations, by stratifying by community). To identify the effects of school quality variables, which are village level attributes, we control for county fixed effects or stratify by county. Dropping these variables, we include village fixed effects or stratify by village. When possible, we also control for village attributes by stratifying by households or by implementing household fixed effects, although the effective sample is much reduced because it includes only households with more than one child. In some cases, especially for the years of schooling hazard, the effective sample for village and household stratification is too small for estimation. Also, when employing county fixed effects or stratification by county we allow for error correlation (or clustering) within villages, adjusting reported standard errors appropriately. When employing village fixed effects or stratification by village, we allow for clustering by household.

6. Results

6.1. Enrollment

The enrollment rate for children in our sample is 78.4 percent — 81.8 percent for boys and 74.4 percent for girls (Table 2). Using a sample of 8000 households in 19 provinces, Knight and Song (2000) calculate a rural enrollment rate of 91 percent for children aged 7-12 and 87 percent for children aged 13-15. In our poor county sample, the enrollment rates for the same age groups are 92.1 percent and 71.2 percent, respectively. The lower enrollment of 13-15 year olds in poor areas is striking considering the fact that children in poor areas tend to enroll at older ages, so that many 13-15 year olds are not in junior secondary school but rather in primary school where enrollment rates tend to be higher. Nationally, the percentage of poor counties with junior secondary schooling enrollment above 85 percent is only 40 percent, compared to 70 percent in all counties (World Bank, 2000).

[Table 2 HERE]

Figure 1 summarizes the enrollment status of children in the sample. School dropouts comprise 12.9 percent of the sample, and form a sample of students who have completed their educations, assuming that they do not subsequently return to school. Of these, 49.1 percent do not reach junior secondary school, 23.6 percent drop out during junior secondary school, and the remaining 27.3 percent withdraw just after completing junior secondary school, oftentimes involuntarily because they cannot pass senior secondary entrance exams. The mean number of grades completed among dropouts is 5.89. In our sample, male dropouts complete 6.5 years of schooling while female dropouts complete 5.3 years. The drop out rate for girls relative to boys is particularly high in the first three years of primary school (Figure 3). Students begin dropping out in earnest as early as age 12. Among 16-year-olds who had ever enrolled in school, 62.2 percent had dropped out (Figure 4).

[Figure 3 HERE]

Parents of dropouts were asked to select from a list of reasons for withdrawing their children from school, and they appear to be less willing to pay for the education of girls. For primary school dropouts, inability to pay high fees, the most frequent response, led to the drop out decision for 47 percent of girls, but only 33 percent of boys, while for junior secondary school dropouts, high fees were cited for half of the girls but only 8 percent of the boys.[14] This is consistent with a higher price elasticity of education for girls, which is found in many developing countries (World Bank, 2000).

[Figure 4 HERE]

Because the factors affecting the decision to continue schooling in primary and junior secondary schools may be different, we look separately at the number of years of schooling for children who ever enrolled at each level. Table 3 presents results from Cox proportional hazard models of the likelihood of stopping schooling at each level.[15] We stratify by county and by village to control for regional and community-level unobserved heterogeneity.[16]

[Table 3 HERE]

Conditional on having remained in school until the current time, the probability that poor and credit constrained children will drop out of primary school is three times that of other children. Thus, it is not surprising that just 6.9 percent of those who had ever enrolled in junior secondary school are poor and credit constrained, while 13.9 percent of primary school enrollees are. Children from poor and credit constrained households who enroll in junior secondary school are much less likely to drop out, perhaps the result of a selection process in which only top students or children of particularly supportive parents remain by junior secondary school. Higher wealth (expenditures per capita) reduces the likelihood of dropping out from primary school, but the coefficient is not statistically significant. The number of siblings reduces the likelihood of dropping out, again suggesting that siblings either substitute for each other's household labor contributions or provide complementarities through cost saving or improved learning.

Variables reflecting intra-household decision-making also affect the duration of schooling. For each additional year of a father's education, the probability of his child dropping out of school falls by 12-14 percent. Also, children of empowered women are much less likely to drop out of primary school. The degree to which women's empowerment plays a role is smaller for girls (a finding consistent with Knight and Song, 2000), although the coefficient is not statistically significant. Finally, boys are more likely to drop out of junior secondary school. There is no statistically significant difference in the probability of boys and girls dropping out of primary school.

The probability of dropping out falls as school fees and distance to school increase. This finding is consistent with higher school fees being charged by higher quality schools. The coefficient is only significant at the primary level, suggesting that junior secondary school fees may be sufficiently high to be a deterrent to enrollment. The inverse relationship between distance to school and the probability of dropping out at the primary level is unexpected if distance increases the costs of schooling because of children's opportunity cost of time. However, the negative coefficient is consistent with a low opportunity cost for primary school students and a positive correlation between distance and school quality, as suggested above. When village strata are included, the coefficient on distance becomes much smaller in magnitude and is no longer statistically significant, which is consistent with our school quality story since identification is coming from within-village differences (where there are no quality effects) rather than differences between villages. The quality of infrastructure also enters the primary school decision in an intuitive way: as the percentage of rainproof classrooms rises, the likelihood of dropping out falls significantly.[17] Finally, the percentage of teachers with post-secondary education positively affects the probability of dropping out at the primary level. We hypothesize that this result stems from teacher education being negatively correlated with teacher experience. The opposite is true for middle schools, although the coefficient is not quite significant at the 90 percent confidence level. Unfortunately, we do not have data on the experience of individual teachers.[18]

6.2. Examination scores

Table 4 presents estimation results for the determinants of the standardized average examination scores of students who were enrolled the previous semester. We incorporate county-grade fixed effects, village-grade fixed effects, and household fixed effects in separate specifications. Expenditures per capita has a robustly positive impact on test scores (a 10 percent increase in expenditures increases test scores by 0.05 standard deviations), suggesting that poverty may reduce human capital accumulation even when enrollment rates are high. The poor and credit constrained dummy, however, is statistically insignificant. Also, children with older siblings have significantly higher test scores than their peers, possibly because they receive help from siblings or because older children substitute their own household labor for the enrolled child's.

[Table 4 HERE]

Controlling for other covariates, junior secondary school girls outperform boys by 0.2-0.7 standard deviations, but the test score gender gap at the primary school level is not significantly different from zero. The performance gender gap is consistent with a gender selection story in which academically weak girls drop out in primary school but academically weak boys do not. The estimated junior secondary school gender difference (and statistical significance) falls as one moves from the county fixed effects specification to the village and household fixed effects specifications, perhaps suggesting that selection effects are greater in areas of poorer average performance, so that a greater share of girls with test scores are in the schools with better performance. Alternative explanations for higher female junior secondary school test scores are that girls study harder than boys or that girls have fewer distractions or other responsibilities that compete for their time. While the latter explanations cannot be ruled out, we find them unlikely.

There is some evidence that parental education and women's empowerment have a negative effect on test scores, although the coefficients are not significant in all specifications. The negative effect could reflect a higher opportunity cost of time for educated parents and empowered women or a selection story in which such parents keep children in school longer, even when their children are academically weak. The negative effect of women's empowerment is significant for girls only, providing weak support for greater gender bias by mothers than fathers. Finally, and somewhat surprisingly, the school quality variables do not enter these regressions significantly, suggesting that there is no effect of school quality on learning, that our measures of school quality do not capture important school quality attributes, or that our small sample of 40 primary and 37 junior secondary schools does not have enough variation for identification.

6.3. Grade promotion

Of students who have ever been enrolled, 30.4 percent have been held back at least one year (Table 2). The mean number of years held back among those ever held back is 1.30, statistically identical for boys and girls. Of those who are ever held back, 74.4 percent are held back just one year. Ministry of Education (2000) observes that the percent of children held back is relatively high in first grade, but falls in every subsequent year. The propensity to be held back varies considerably by province; in Sichuan, for example, just 13 percent have ever been held back, while 47 percent have been held back in Shaanxi. Boys are more likely than girls to have been held back (35 and 25 percent, respectively), which may be because boys have poorer study habits or because boys go farther in their education and so have more chances to be held back.

Patterns in the percent of students ever held back and the average number of years held back among those ever held back provide support for a story in which poorly performing girls are more likely to drop out in primary school. If all children stay in school even if they are held back, or if children drop out for reasons uncorrelated with being held back, the percentage of children ever held back should increase with age since more time in school increases the number of chances of being held back. However, in our sample, the percentage of students ever held back falls with age for girls but not for boys, direct evidence that girls who are held back are relatively more likely to drop out. Also, as seen in Figure 5, the average number of years held back among those ever held back increases with age for boys but not for girls. Finally, comparing the performance of boys and girls among dropouts and non-dropouts, we find that the ratio of the share of boys ever held back to the share of girls ever held back is 0.8 for dropouts but 1.4 for non-dropouts. In other words, among those in school, boys are more likely to have been held back, but among dropouts, girls are more likely to have been held back. These patterns provide evidence for a differential selection story, and unlike the test score results, they have no plausible alternative explanation. King, Orazem, and Paterno (1999) show that in the Philippines, too, promotion is a much stronger predictor of continued enrollment for girls than for boys.

[Figure 5 HERE]

Table 5 presents results for a conditional logit model for ever having been held back. Odds ratios and coefficients are reported for specifications including county, village, and household fixed effects. Dummies for the number of grades completed have been included to control for the number of opportunities to be held back.

[Table 5 HERE]

Consistent with the gender selection story in which academically weak boys stay in school longer than academically weak girls, boys are more likely to have been held back than girls, particularly for those in junior secondary school. The economic and school quality variables do not significantly affect the likelihood of ever being held back. Expenditures per capita enters positively in the county fixed effects specification but not in the village fixed effects specification. One of the most significant factors affecting whether children are held back is the age of enrollment. Kids who enroll later are less likely to have ever been held back, consistent with our expectation that many of the children who are held back are those who enter school earlier and are held back in first grade. If promotion after first grade is relatively automatic, it may explain why many variables do not robustly explain whether kids are ever held back.

7. Conclusion

In concluding, we attempt to integrate the important results above to draw broader inferences about the importance of low wealth and credit constraints, intra-household decision-making, and school quality on educational attainment in poor areas.

Poverty significantly affects both educational investments and learning. Controlling for expenditures per capita, children from households that are both poor and credit constrained are three times as likely to drop out of school. Thus, for some of the poor, the lack of available funds is a major obstacle to financing educational investments. However, being poor and credit constrained does not significantly affect learning in school (as measured by test scores or being held back), suggesting that the inability to finance educational expenditures does not hurt children's performance in school . There is weak evidence that wealth, measured by expenditures per capita, affects years of schooling independently of whether one is poor and credit constrained, and strong evidence that it affects test scores; a ten percent increase in expenditures increases test scores by 0.05 standard deviations. In addition to pointing out the importance of credit constraints in poor areas, our results show that even when households are not credit constrained, children from wealthier households have an advantage in school performance. Thus, even with high enrollments, poverty still may be an important issue in educational attainment because of its direct effects on learning. The different results for the wealth and credit constraint variables also highlight the value of using more direct measures of credit constraints when evaluating wealth effects on educational attainment.

We find strong evidence of a gender selection story in which poorly performing girls drop out in primary school while boys do not begin to drop out in earnest until junior secondary school. The relative likelihood of having ever been held back is greater for girls among dropouts but greater for boys among those in school. The average age of enrollment in primary school is younger for boys but the average age of enrollment in junior secondary school is younger for girls. These patterns suggest that girls that are held back are more likely to drop out than boys that are held back. In addition, girls score higher on tests in junior secondary school, suggesting a weeding out of poorly performing girls in primary school. The clear gender bias in educational investments may be due to lower returns to education for girls, the lower selfish returns to parents from investing in girls that will marry into other families, or from parental preferences that favor sons. Further research that more convincingly distinguishing among competing explanations for gender bias should receive high priority.

With regard to women's empowerment, the strongest effects occur for years of schooling. The coefficient estimates suggest that the likelihood of dropping out of primary school falls dramatically when women have a greater say in enrollment decisions (but not quite statistically significant with village strata), and the probability of dropping out of junior secondary school falls dramatically for sons. These results imply that women value education more than men, and that if anything they favor sons more than daughters, relative to men. Whether children were ever held back is also significantly influenced by women's empowerment; the empowerment of less educated women has a significantly stronger negative effect on the likelihood of children being held back than the empowerment of better-educated women.

Father's education has a much greater influence on educational investment decisions than mother's education. An additional year of father's education reduces the likelihood of dropping out by 12-14 percent and reduces the likelihood that the child was ever held back by 14 percent. These positive effects may reflect preferences associated with higher education or higher returns to education if children have more educated fathers. This might be true if children of educated fathers have greater ability, if a father's education serves as a complementary input to children's learning, or if more educated parents have better social and professional networks that increase future labor market opportunities. The relative unimportance of mother's education may be partly due to the very low average educational level of women in the study areas.

We find evidence that the presence of siblings reduces the likelihood of dropping out, especially if siblings are older, and that children with older siblings score higher on exams. These findings suggest that children benefit from having siblings, and that younger siblings in particular benefit from the presence of older siblings. Siblings can increase the desirability of educational investments by substituting for each other's labor contributions to the household, economizing on costs, or helping each other with schoolwork.

Finally, our measures of school quality do not appear to affect learning in school (test scores, ever having been held back), but they do have some effect on the years of schooling. Higher school fees and distance to school, each of which may be a proxy for quality, result in lower probabilities of dropping out of primary school. The percentage of classrooms which are rainproof and the percentage of teachers with post-secondary education significantly impact the number of years of primary schooling (although the latter has an unexpected sign which we attribute to a negative correlation between education and experience). Thus, while school attributes do affect educational investment decisions, our estimates do not pick up direct effects on learning. This may be because there are no effects, because our particular measures of school quality do not adequately measure important school attributes in the Chinese context, or because our small sample of schools lacks the variation necessary to detect the effects of differences in school quality. Future research using better measures will be of great interest.

Acknowledgements

We thank the Ford and Luce Foundations for supporting the field surveys in China, and Emily Hannum, David Lam, Lee Lillard, Rohini Somanathan, Sangui Wang, Robert Willis, seminar participants at the Harvard Graduate School of Education, the University of Michigan Department of Economics, the 2000 NEUDC Conference at Cornell University, the Institute of Population Studies (Chinese Academy of Social Sciences), and a Workshop on Poverty Policies in China held by the China Poverty Research Association in Beijing, as well as two anonymous referees for helpful comments.

Notes

[1] By modeling the parents' joint payoff, we abstract from possibly different payoffs to fathers and mothers. We do allow for different parental attitudes toward the welfare of children, however.

[2] We omit consumption in order to simplify; minimum consumption requirements would reduce further the maximum educational investment of credit constrained households.

[3] We include sibling effects to capture competition for resources even though our model includes only one child. It is straightforward to adjust the model to allow for multiple children. Presence of siblings might also affect the expected future contributions from children.

[4] In a systematic analysis using panel data, Glewwe and Jacoby (2000) find wealth effects for Vietnam; see also Filmer (2000) for cross-country evidence, Jacoby (1994) for evidence of the importance of credit constraints in Peru, and Behrman and Knowles (1999) for a review of the issues. Schultz (2000) and World Bank (2000) show that tuition subsidies increased enrollment in Mexico and other developing countries. Foster and Rosenzweig (1996), however, find no wealth effects for India.

[5] These cutoffs are, of course, arbitrary. However, a simple interaction term between expenditures per capita and credit constraints would miss an important nonlinearity. We tried different cutoff values and chose the highest values for which the effects were significant. Lower cutoffs produced similarly significant results.

[6] There is mixed evidence on whether men and women favor sons and daughters differently (World Bank, 2000). For example, Lillard and Willis (1994) find that mothers' education has a greater impact on daughters' education and that fathers' education has a greater impact on sons' education in Taiwan, but Quisumbing and Maluccio (1999) find the opposite in South Africa.

[7] Case and Deaton (1999) find significant effects of student-teacher ratios in South Africa. Birdsall (1985) finds strong effects of teacher's education in Brazil. Hanushek (1995) concludes from a review of previous studies that teacher training deserves greater support.

[8] 5.5 percent of our sample had enrolled before age six. Interviews suggest in most cases such children have siblings already attending school.

[9] We estimated age of enrollment hazards in which we stratified by county, village, and household. We found that the number of older siblings positively impacts the probability of enrollment, perhaps because an older siblings can

accompany younger ones to school and can provide hand-me-downs that reduce the cost of schooling. Having a higher birth order reduces the probability of primary school enrollment, perhaps because families have accumulated less wealth when their first children are born or because older children contribute to the household in ways that make it easier to send younger children to school. Surprisingly, per capita expenditure has no discernable effect on the age of enrollment, and children who are both poor and credit constrained are likely to enroll earlier in primary school. We also find that boys enroll earlier in primary school than girls (50 percent more likely to enroll all things equal), and that an additional year of father's education raises the probability of earlier enrollment by 9 percent. Finally, our school quality variables do not enter the initial enrollment decision significantly.

[10] Using the stratified Cox proportional hazard model, the hazard at time *t* for a child i in community *j* is assumed to equal

$$h_i(t) = h_{0j}(t)e^{\beta 1x1i + \ldots + \beta kxki}$$

[11] For example, a hazard ratio of 1.5 means that the child is 1.5 times more likely to drop out if the independent variable increases by one unit. Thus, hazard ratios greater than one correspond to positive coefficients and hazard ratios less than one correspond to negative coefficients.

[12] In most of the study regions, family planning policy dictates up to two children. Areas of the county in Guizhou may allow up to three children, since it is a minority county. The data show that two thirds of the families in the sample have two or fewer children, and 95 percent of families have three or fewer children. This would appear to be roughly consistent with the expected effects of the family planning policy.

[13] There are no obvious identifying variables for a selection equation. Many researchers have used distance to school and tuition as identifying variables. However, because our estimates find both to have positive effects on enrollment, we believe they are likely to reflect school quality differences, which may affect learning outcomes. Tuition, in particular, seems strongly associated with the wealth of the community and the quality of schools. In addition to our previous explanation about distance as an indicator of school quality, distance also could affect attendance (unobserved), ability to spend more time in school after class, etc., which could have effects on learning outcomes. Without good identifying variables, a Heckman-type selection model can only be identified from the assumption of joint normality of the error terms in the selection and outcome equations. We estimate the selection model with and without distance as an identifying variable, and obtain similar results.

[14] For primary dropouts, the other main reasons cited were child unwillingness to attend school (44 percent), the interpretation of which is unclear, and poor grades (7 percent). For junior secondary dropouts, 35 percent reported poor grades, and 15 percent cited an unwillingness to attend.

[15] We could alternatively use an ordered logit model to estimate the determinants of grade attainment. We feel that years of schooling has a more natural behavioral interpretation since the decision of families is to keep the child in school, not to promote the child to the next grade. Grade attained also conflates the decision to stay in school and performance in school, and one of our goals is to consider these aspects of educational attainment separately. Nonetheless, we did estimate an ordered logit model of grade attainment, and found the important results to be quite similar.

[16] We do not stratify by village in the junior secondary school estimates, nor by household in either the primary or junior secondary school estimates. In each case, there was too little variation within the strata of concern.

[17] The rainproof classrooms variable was omitted from the junior secondary school hazard because it perfectly predicted dropping out in some counties due both to the small number of junior secondary school dropouts in the sample and to the large number of junior secondary schools without any leaking classrooms.

[18] To check the robustness of results to the included variable set, we also estimated the hazard models dropping the age of enrollment, and find no substantial changes. In general, the model is less precisely estimated, and in no cases do variables gain significance when the age of enrollment is excluded.

References

Behrman, J.R., & Knowles, J.C. (1999). Household income and child schooling in Vietnam. *World Bank Economic Review* 13(2), 211-256.

Benjamin, D., Brandt, L., Glewwe, P. & Li, G. (1999). Markets, human capital, and inequality: evidence from rural China. *Mimeo*, Development Research Group, World Bank.

Birdsall, N. (1985). Public inputs and child schooling in Brazil. *Journal of Development Economics* 18(1), 67-86.

Bound, J., Jaeger, D.A., & Baker, R.M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. *Journal of the American Statistical Association* 90(430), 443-450.

Case, A., & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics*. 114(3), 1047-1084.

Connelly, R., & Zheng, Z. (2000). Determinants of primary and middle school enrollments of 10 to 19 year olds in China. *Mimeo*, Department of Economics, Bowdoin College.

Cox, D.R., & Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman Hall.

Filmer, D. (2000). The structure of social disparities in education. *World Bank Policy Research Working Paper No. 2268*. Washington, D.C.: World Bank.

Foster, A.D., & Rosenzweig, M.R. (1996). Technical change and human-capital returns and investments: Evidence from the Green Revolution. *American Economic Review* 86(4), 931-953.

Glewwe, P., Grosh, M., Jacoby, H., & Lockheed, M. (1995). An eclectic approach to estimating the determinants of achievement in Jamaican primary education. *World Bank Economic Review* 9(2), 231-258.

Glewwe, P., & Jacoby, H. (1994). Student achievement and schooling choice in low income countries: Evidence from Ghana. *Journal of Human Resources* 29(3), 841-864.

Glewwe, P., & Jacoby, H. (2000). Economic growth and the demand for education: Is there a wealth effect?. *Paper presented at the New Research on Education in Developing Countries conference*, Stanford University.

Hannum, E. (1998). Educational inequality: Hidden consequences of the reform era in rural China. *Ph.D. Dissertation*, Department of Sociology, University of Michigan.

Hannum, E. & Xie, Y. (1994). Trends in educational gender inequality in China: 1949-1985. *Research in Social Stratification and Mobility* 13, 73-98.

Hanushek, E.A. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer* 10(2), 227-246.

Hanushek, E.A., & Lavy, V. (1994). School quality, achievement bias, and dropout behavior in Egypt. *Living Standards Measurement Study Working Paper, No. 107.* Washington, D.C.: World Bank.

Hossain, S.A. (1996). Making an equitable and efficient education: The Chinese experience. *Mimeo*, revised from *China: Social Sector Expenditure Review, 1996.* Washington, D.C.: World Bank.

Jacoby, H.G. (1994). Borrowing constraints and progress through school: Evidence from Peru. *Review of Economics and Statistics* 76(1), 151-160.

Jamison, D.T., & Lockheed, M.E. (1987). Participation in schooling: Determinants and learning outcomes in Nepal. *Economic Development and Cultural Change* 35(2), 279-306.

Jamison, D.T., & van der Gaag, J. (1987). Education and earnings in the People's Republic of China. *Economics of Education Review* 6(2), 161-166.

Khandker, S.R. (1996). Education achievements and school efficiency in rural Bangladesh. *Discussion Papers, No. 319.* Washington, D.C.: World Bank.

King, E.M., Orazem, P.F. & Paterno, E.M. (1999). Promotion with and without learning: Effects on student dropout. *Working Paper Series on Impact Evaluation of Education Reforms 18*. World Bank, Development Research Group.

Knight, J., & Song, L. (1999). Increasing wage inequality in China: Efficiency versus equity? *University of Oxford Applied Economics Discussion Paper Series, No. 211.* Oxford University, Department of Economics.

Knight, J., & Song, L. (2000). Differences in educational access in rural China. *Mimeo*, Department of Economics, University of Oxford.

Lee, J.M., and Wang, F. (1999). *One Quarter of Humanity: Malthusian Mythology and Chinese Realities*. Harvard: Harvard University Press.

Lillard, L., & Willis, R. (1994). Intergenerational educational mobility: Effects of family and state in Malaysia. *Journal of Human Resources* 29(4),1126-1166.

Ministry of Education, National Commission for UNESCO, People's Republic of China. (2000). *National Report For EFA 2000 Assessment*. Beijing: Ministry of Education.

Parish, W.L., & Willis, R.J. (1993). Daughters, education and family budgets: Taiwan experiences. *Journal of Human Resources* 28(4), 863-898.

Park, A., Rozelle, S., Wong, C., & Ren, C. (1996). Distributional consequences of reforming local public finance in China. *The China Quarterly* 147, 751-778.

Park, A., & Wang, S. (2000). Will credit access help the poor?: Evidence from China. *Mimeo*, Department of Economics, University of Michigan.

Quisumbing, A.R., & Maluccio, J. (1999). Intrahousehold allocation and gender relations: New empirical evidence. *Policy Research Report on Gender and Development Working Paper 2*. Washington, D.C.: World Bank.

Schultz, T.P. (2000). School subsidies for the poor: Evaluating A Mexican strategy for reducing poverty. *Paper presented at the New Research on Education in Developing Countries conference*, Stanford University.

Tsang, M.C. (1994). Costs of Education in China: Issues of Resource Mobilization, Equality, Equity, and Efficiency. *Education Economics* 2(3), 287-312.

Tsang, M.C. (1996). Financial reform of basic education in China. *Economics of Education Review* 15(4), 423-444.

West, L. (1996). Provision of public services in rural PRC. In: Christine Wong, ed., *Financing Local Government in the People's Republic of China* (pp. 213-282). Hong Kong: Oxford University Press for the Asian Development Bank.

World Bank. (1999). *Strategic Goals for Chinese Education in the 21st Century*. Washington D.C.: World Bank.

World Bank. (2000). *Engendering Development*. Washington D.C.: World Bank.

Figure 1
Enrollment status

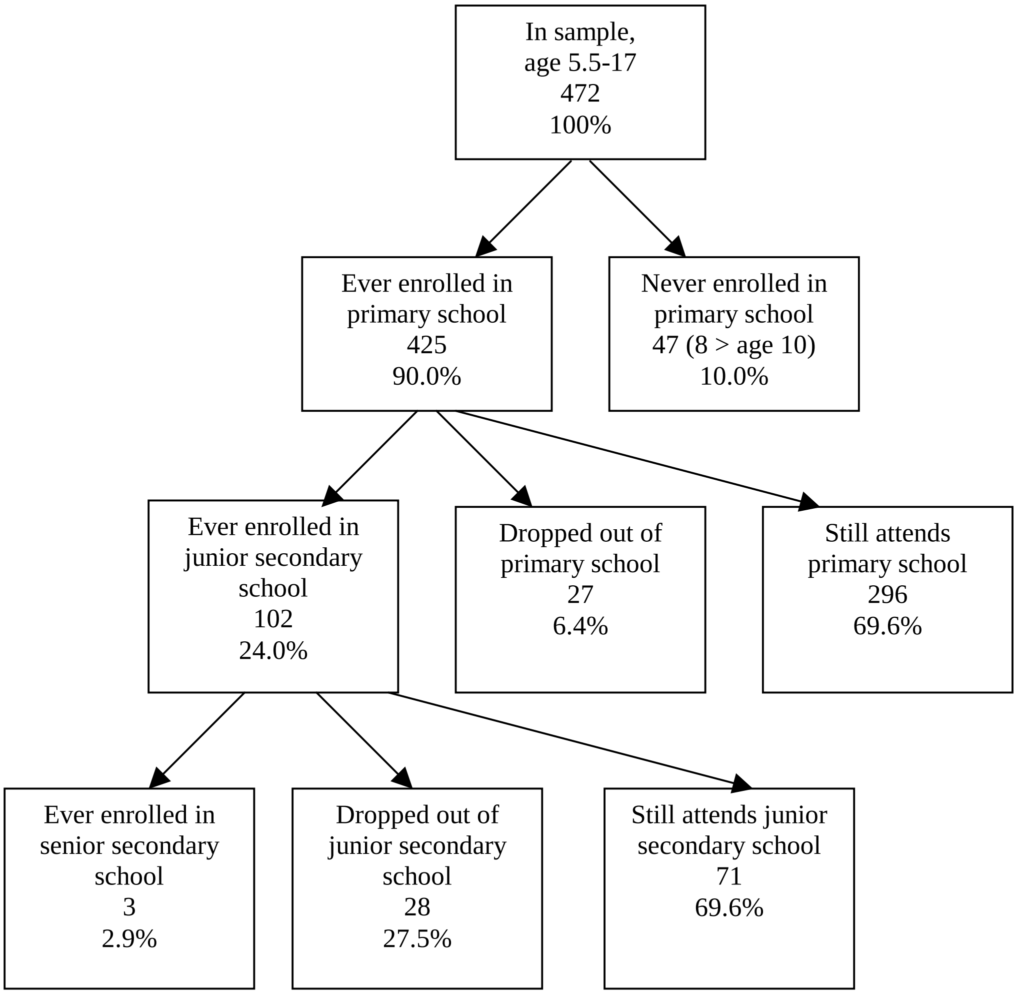

Figure 2A
Age of primary school enrollment by sex

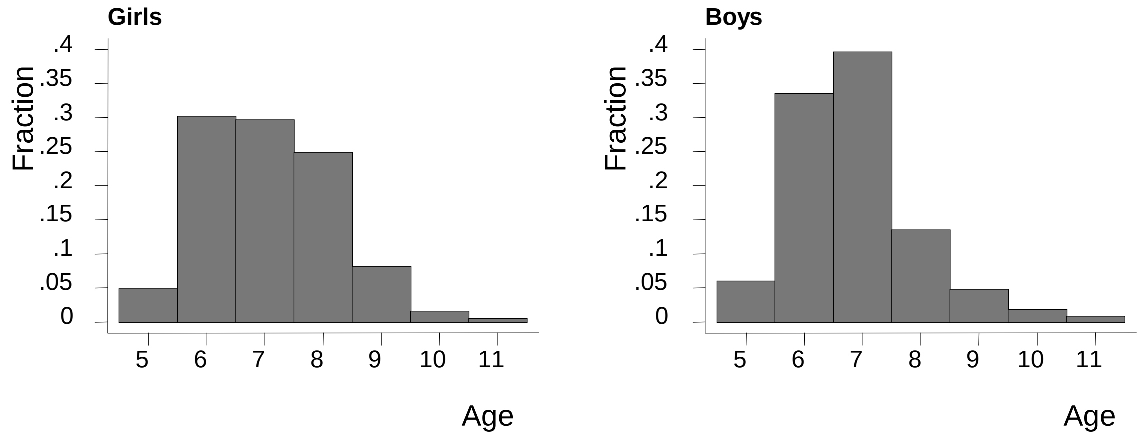

Figure 2B
Age of junior secondary school enrollment by sex

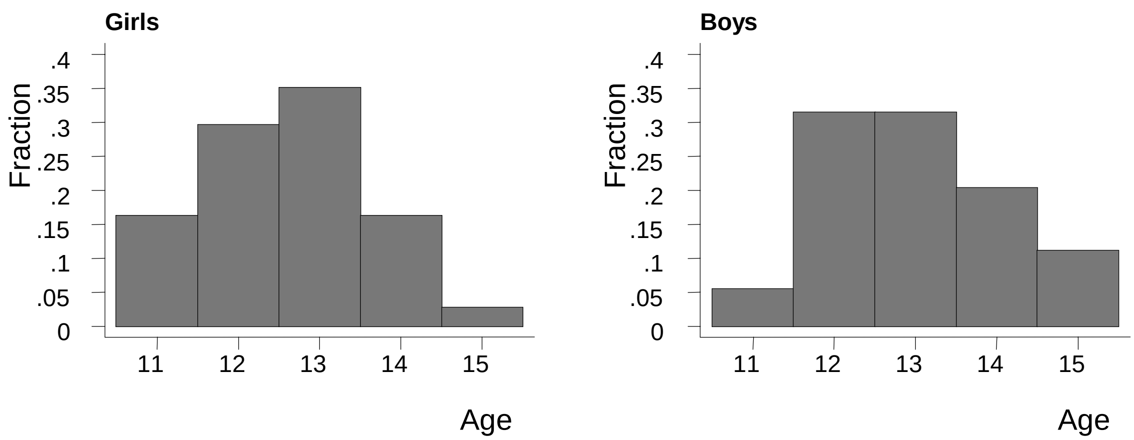

Figure 3
Years of schooling among dropouts

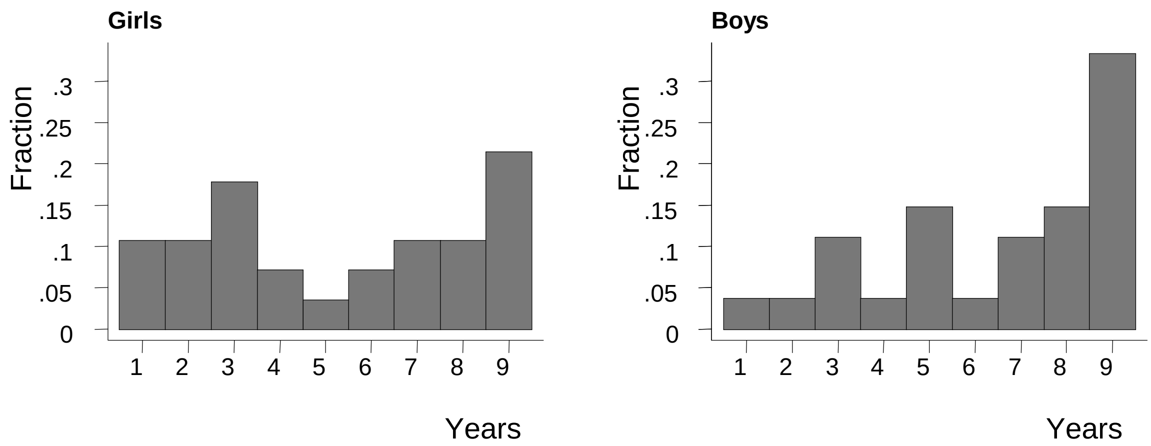

Figure 4
Enrollment rate

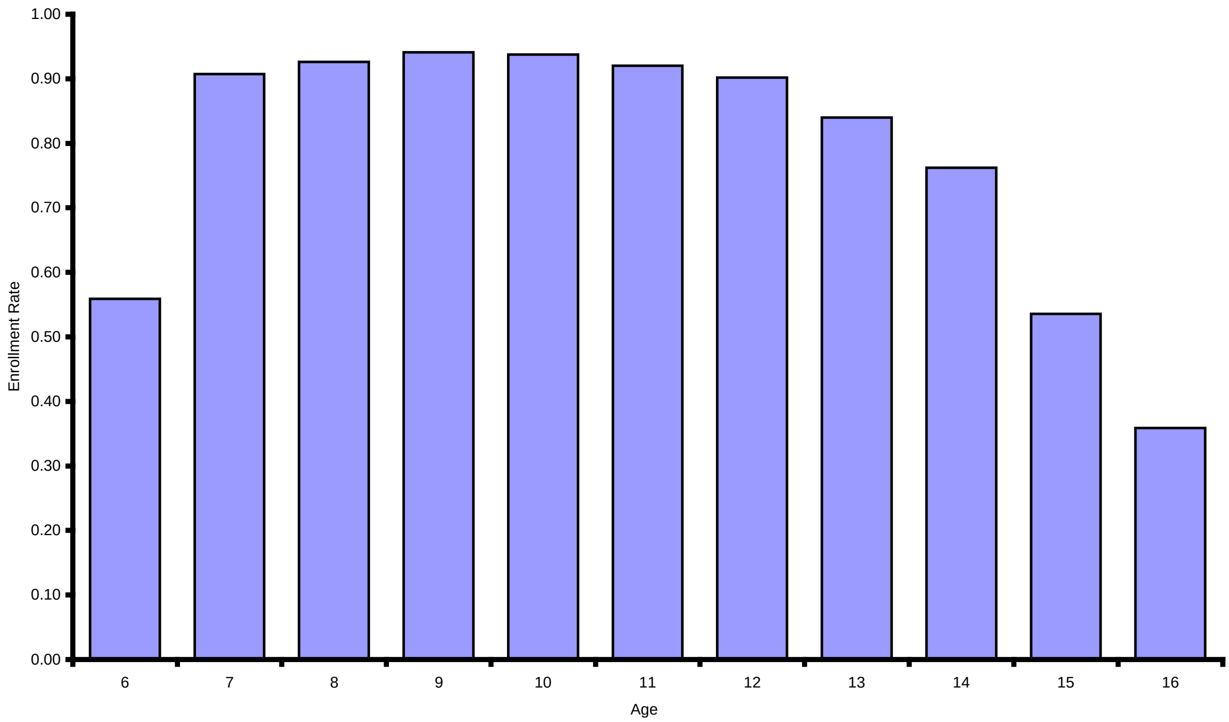

Figure 5
Mean number of years held back (among those ever held back)

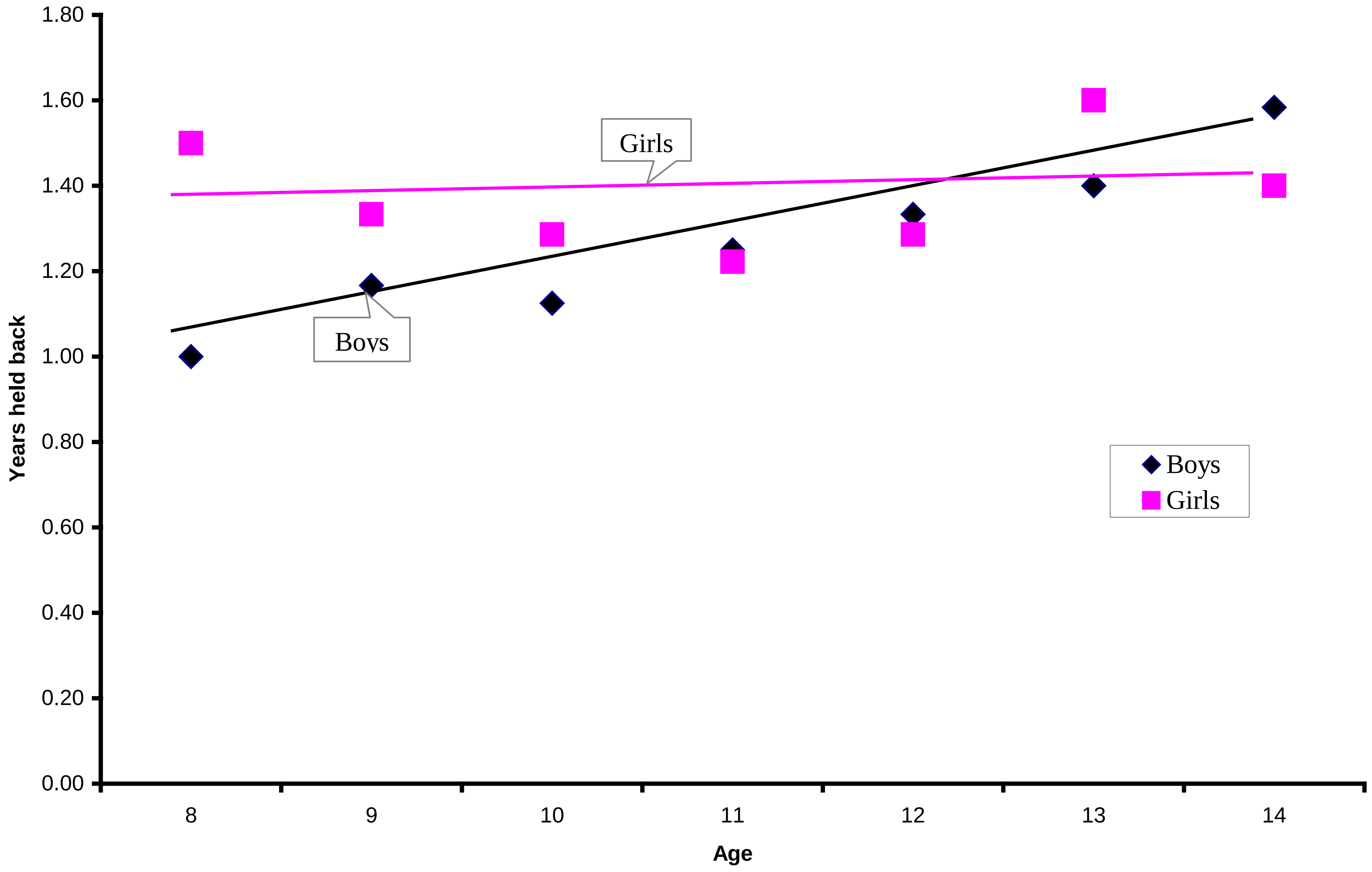

Table 1
Summary statistics for sample households and schools

	Variable	Unit	Obs	All Mean	Std dev	Means Henan	Jiangxi	Sichuan	Guizhou	Shaanxi	Gansu
Households	father's education	years	261	7.31	3.40	8.00	7.24	7.58	4.92	8.69	7.25
	mother's education	years	252	3.35	3.40	3.31	3.35	5.66	1.00	4.28	2.66
	women's empowerment	%	262	0.26	0.31	0.28	0.12	0.29	0.29	0.36	0.24
	number of children	#	262	2.15	0.84	2.28	1.93	1.98	2.12	2.16	2.44
	per capita expenditure	yuan	262	1133.54	651.97	857.10	1682.67	1718.75	971.18	687.60	811.81
	credit limit	yuan	262	4643.41	6596.45	6513.85	5000.29	3866.80	1685.78	6287.22	3468.44
	poor and credit constrained	1/0	262	0.14	0.35	0.18	0.02	0.00	0.25	0.19	0.22
	distance to primary school	km	195	1.22	(1.20)	0.54	1.20	1.56	2.34	0.62	1.24
	distance to junior secondary school	km	71	3.55	(4.87)	4.27	2.78	2.08	2.25	2.46	4.45
primary school	school fees	yuan	40	100.92	(47.01)	73.86	120.42	169.66	42.30	113.28	64.64
	student-teacher ratio	#	40	28.15	(9.49)	31.56	24.21	35.06	22.20	26.27	20.61
	teachers with post-secondary education	%	40	0.54	(0.34)	0.52	0.61	0.55	0.28	0.86	0.33
	rainproof classrooms	%	40	0.78	(0.51)	0.66	1.00	0.82	0.74	0.80	0.70
junior sec. School	school fees	yuan	37	317.80	(122.98)	295.68	304.95	383.90	216.86	400.25	294.83
	student-teacher ratio	#	37	15.10	(6.63)	14.10	18.78	18.64	11.66	11.39	11.37
	teachers with post-secondary education	%	37	0.88	(0.20)	0.91	0.95	0.85	0.79	0.93	0.83
	rainproof classrooms	%	36	0.90	(0.18)	0.95	1.00	0.94	1.00	0.80	0.68

Table 2
Educational attainment indicators

		Students, age 5.5-17			Boys			Girls		
Variable	Unit	Obs	Mean	Std dev	Obs	Mean	Std dev	Obs	Mean	Std dev
current enrollment rate	%	472	0.78	(0.34)	253	0.82	(0.39)	219	0.74	(0.44)
age of primary enrollment	years	418	7.46	(1.13)	231	7.34	(1.07)	187	7.60	(1.17)
age of junior secondary enrollment	years	91	13.38	(1.04)	54	13.52	(1.08)	37	13.17	(0.94)
ever held back	%	425	0.30	(0.46)	234	0.35	(0.48)	191	0.25	(0.43)
years held back, if held back	years	129	1.30	(0.55)	81	1.30	(0.56)	48	1.31	(0.55)
highest grade completed among dropouts	#	55	5.89	(2.79)	26	6.51	(2.58)	28	5.28	(2.92)
average years per grade	years	336	1.17	(0.37)	234	1.21	(0.27)	152	1.14	(0.37)

Table 3
Years of schooling (Cox proportional hazard)

		Primary school 1			Primary school 2			Junior secondary school		
Variable	Unit	Hazard ratio	Coef.	Std. error	Hazard ratio	Coef.	Std. error	Hazard ratio	Coef.	Std. error
male	1/0	1.148	0.138	(0.601)	0.599	-0.513	(0.447)	3.102**	1.132	(0.500)
age of enrollment	years	1.430*	0.358	(0.207)	1.090	0.086	(0.139)	0.477*	-0.740	(0.440)
younger siblings	#	0.654	-0.424	(0.313)	0.576	-0.551	(0.417)	0.577	-0.549	(0.409)
older siblings	#	0.147**	-1.917	(0.938)	0.138*	-1.980	(1.108)	0.617	-0.483	(0.784)
father's education	years	0.881	-0.127	(0.079)	0.875	-0.134	(0.088)	0.855	-0.157	(0.108)
mother's education	years	0.975	-0.025	(0.154)	0.960	-0.041	(0.120)	0.921	-0.082	(0.083)
women's empowerment	percent	0.087**	-2.441	(1.103)	0.209	-1.566	(1.212)	3.220	1.170	(2.153)
women's empowerment * male	interaction	0.558	-0.583	(1.436)	0.366	-1.005	(1.384)	0.055**	-2.898	(1.394)
women's empowerment * mother's education	interaction	1.057	0.055	(0.436)	1.102	0.097	(0.301)	1.157	0.145	(0.321)
log expenditure per capita	log yuan	0.250	-1.386	(1.766)	0.833	-0.183	(0.932)	0.577	-0.550	(0.831)
poor and credit constrained	1/0	4.891**	1.587	(0.642)	3.047	1.114	(0.782)	0.244*	-1.412	(0.854)
distance to school	km	0.475**	-0.745	(0.371)	0.845	-0.169	(0.114)	0.901	-0.105	(0.077)
log school fees	log yuan	0.435*	-0.832	(0.496)				0.993	-0.007	(0.629)
student-teacher ratio	#	0.968	-0.032	(0.034)				0.998	-0.002	(0.044)
rainproof classrooms	%	0.047**	-3.067	(0.970)						
teachers with post-secondary education	%	4.164*	1.427	(0.758)				0.131	-2.034	(1.256)
county strata			X						X	
village strata						X				
N			373			188			79	
Prob>chi2			0.0000			0.0012			0.0000	

* significant at the 90 percent confidence level
** significant at the 95 percent confidence level

Table 4
Standardized examination scores (ordinary least squares)

Variable	Unit	1 Coeff.	1 Std. error	2 Coeff.	2 Std. error	3 Coeff.	3 Std. Error
male	1/0	-0.668**	(0.296)	-0.443*	(0.268)	-0.218	(0.321)
male * primary	interaction	0.646**	(0.281)	0.342	(0.226)	0.099	(0.337)
age of enrollment	years	-0.020	(0.081)	0.016	(0.082)	-0.032	(0.102)
younger siblings	#	0.152	(0.129)	0.194	(0.138)	-0.117	(0.106)
older siblings	#	0.268**	(0.113)	0.287**	(0.139)		
father's education	years	-0.021	(0.024)	-0.045*	(0.025)		
mother's education	years	-0.018	(0.025)	-0.007	(0.032)		
women's empowerment	%	-0.635**	(0.308)	-0.693	(0.449)		
women's empowerment * male	interaction	0.452	(0.418)	0.619	(0.407)		
women's empowerment * mother's education	interaction	0.080	(0.059)	0.068	(0.066)		
log expenditure per capita	log yuan	0.535**	(0.176)	0.544**	(0.237)		
poor and credit constrained	1/0	0.326	(0.218)	0.371	(0.247)		
distance to school	km	0.001	(0.016)	-0.007	(0.016)		
log school fees	log yuan	0.072	(0.158)				
student-teacher ratio	#	-0.015	(0.019)				
rainproof classrooms	%	-0.319	(0.828)				
teachers with post-secondary education	%	-0.316	(0.621)				
log school fees * primary	interaction	-0.336*	(0.187)				
student-teacher ratio * primary	interaction	0.012	(0.017)				
rainproof classrooms * primary	interaction	0.716	(0.870)				
teachers with post-secondary education * primary	interaction	0.726	(0.622)				
constant		-2.813*	(1.503)	-3.663**	(1.731)	0.382	(0.742)
county fixed effects		X					
village fixed effects				X			
household fixed effects						X	
N		260		271		176	
R-squared		0.11		0.25		0.71	

* significant at the 90 percent confidence level
** significant at the 95 percent confidence level
Notes: Dependent variable is standard deviations from mean test score for the same county-grade. The variable "primary" is a dummy variable for whether the child is currently enrolled in primary school.

Table 5
Ever held back (conditional logit)

			1			2			3	
Variable	Unit	Odds ratio	Coef.	Std. error	Odds ratio	Coef.	Std. error	Odds ratio	Coef.	Std. error
Male	1/0	3.971**	1.379	(0.653)	3.804*	1.336	(0.709)	2.192	0.785	(0.939)
male * primary	interaction	0.468	-0.759	(0.663)	0.461	-0.774	(0.725)	0.953	-0.048	(0.978)
age of enrollment	years	0.701**	-0.355	(0.133)	0.693**	-0.367	(0.158)	0.320**	-1.141	(0.385)
younger siblings	#	1.381	0.323	(0.214)	1.095	0.091	(0.245)	2.540**	0.932	(0.325)
older siblings	#	1.081	0.078	(0.218)	0.764	-0.269	(0.264)			
father's education	years	0.864**	-0.146	(0.048)	0.855**	-0.157	(0.056)			
mother's education	years	0.921	-0.083	(0.055)	0.902	-0.103	(0.065)			
women's empowerment	%	0.788	-0.238	(0.738)	0.511	-0.672	(0.807)			
women's empowerment * male	interaction	0.685	-0.378	(0.866)	0.961	-0.039	(0.909)			
women's empowerment * mother's education	interaction	1.183	0.168	(0.119)	1.283*	0.249	(0.141)			
log expenditure per capita	log yuan	1.988	0.687	(0.420)	1.329	0.284	(0.468)			
poor and credit constrained	1/0	0.873	-0.136	(0.379)	1.459	0.378	(0.446)			
distance to school	km	0.948	-0.053	(0.058)	0.962					
log primary school fees	log yuan	0.628	-0.466	(0.582)						
primary student-teacher ratio	#	0.990	-0.011	(0.021)						
primary rainproof classrooms	%	0.528	-0.639	(0.694)						
primary teachers with post-secondary education	%	0.689	-0.372	(0.484)						
completed second grade	1/0	1.084	0.080	(0.401)	1.349	-0.039	(0.064)			
completed third grade	1/0	1.338	0.291	(0.487)	0.670	0.300	(0.412)			
completed fourth grade	1/0	1.142	0.133	(0.484)	1.360	-0.400	(0.526)			
completed fifth grade	1/0	0.680	-0.386	(0.658)	0.915	0.307	(0.535)			
completed sixth grade	1/0	1.611	0.477	(0.824)	2.171	-0.089	(0.808)			
completed seventh grade	1/0	0.450	-0.798	(0.657)	0.386	0.775	(0.928)			
county fixed effects			X							
village fixed effects						X				
household fixed effects									X	
N			359			328			113	
Prob>chi2			0.0057			0.1243			0.2617	

* significant at the 90 percent confidence level
** significant at the 95 percent confidence level
Notes: The dependent variable equals one if the student was ever held back and zero otherwise. Sample includes all students who were ever enrolled. The variable "primary" is a dummy variable for whether the child is currently enrolled in primary school or dropped out before entering junior secondary school.