Education and the poverty trap in rural China

Knight, John; Shi, Li; Quheng, Deng

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

RECOUP
RESEARCH CONSORTIUM ON
EDUCATIONAL OUTCOMES & POVERTY

RECOUP Working Paper 11

Education and the Poverty Trap in Rural China

John Knight, Li Shi and Deng Quheng
University of Oxford
December 2007

DFID Department for International Development

UNIVERSITY OF CAMBRIDGE

RECOUP Working Paper No. 11

Education and the Poverty Trap in Rural China

John Knight, Li Shi and Deng Quheng

University of Oxford

Abstract

This is an ambitious attempt to view the relationships involving education and income as forming a system, and one that can generate a poverty trap. The setting is rural China, and the data are from a national household survey for 2002, designed with research hypotheses in mind. Enrolment is high in rural China by comparison with most poor rural societies, but the quality of education varies greatly. There are three main strands to the paper. One examines the determinants of enrolment, and finds that poverty has an adverse effect on both the quality and quantity of education - so contributing to a poverty trap. The second examines the effects of education. It shows how and why the returns to education vary according to household and community income – so also contributing to a poverty trap. The third strand brings no fewer than 17 estimated relationships together as a system, and poses the question: can education break the vicious circle of poverty? The implications for poverty analysis and for educational policy are considered.

Correspondence: John Knight, Department of Economics, Manor Road Building, Oxford OX1 3UQ, john.knight@economics.ox.ac.uk

JEL Classification: I20, I30, I32

Key Words: Education; school enrolment; quality of education; poverty; poverty trap; vicious circle; virtuous circle; cumulative causation; credit constraint; China

Acknowledgements: The research was conducted while Li Shi and Deng Quheng were visiting the Department of Economics, University of Oxford, from Beijing Normal University and the Chinese Academy of Social Sciences Graduate School of Economics respectively. We are grateful to Nina Fenton for her research assistance, to Chris Colclough for his comments, and to the UK Department for International Development (DfID) for financial support through John Knight's membership of the *Research Consortium on Educational Outcomes and Poverty (RECOUP)*, which is based partly at the Centre for the Study of African Economies in the Department of Economics, University of Oxford.

Outline

1. Introduction

2. Background

3. Literature review
   3.1 The determinants of enrolment
   3.2 The effects of education

4. Theory and hypotheses: the vicious circle of poverty
   4.1 Poverty traps
   4.2 The interdependent system

5. Descriptive information

6. The determinants of enrolment
   6.1 Explaining middle school drop-out
   6.2 Explaining high school continuation

7. Further links in the circle of deprivation
   7.1 Explaining the influence of community
   7.2 Explaining school performance
   7.3Explaining the determinants of educational quality

8. The effects of education
   8.1 Education and labour allocation between farm and non-farm activities
   8.2 The economic benefits of education in farm and non-farm activities
   8.3 Simulating the effects of raising educational levels
   8.4 Explaining differential returns to education
   8.5 Education, poverty and health
   8.6 Education, poverty and innovation
   8.7 Education, poverty and subjective well-being

9. Summary
9.1 The seventeen relationships
9.2 Can education break the vicious circle of poverty?

10. Conclusions

1. Introduction

This research project falls within the DfID-funded *Research Consortium on Educational Outcomes and Poverty (RECOUP).* It exploits an existing data set, which was partly designed with our research questions in mind. It can be seen as a prototype for research on the four countries on which the Programme is concentrating: India, Pakistan, Kenya and Ghana. The most interesting comparative study, at a later stage of the Programme, would be of China with its giant counterpart, India: rural China has much higher educational attainment and also enrolment than does rural India, and an analysis of the reasons for, and consequences of, this difference would be instructive.

China provides an appropriate case study on the relationships between education and poverty. By comparison with the rural areas of the four countries that are being examined in depth within the programme, rural China provides better educational access, certainly at the primary level, probably at the secondary level, and possibly at the tertiary level. However, there is a wide variation in both the quantity and quality of educational provision, by province, by county and by village. This variation reflects the decentralised system by which public education is funded. Poor areas are at a disadvantage. Moreover, if the general marketisation of the Chinese economy has meant that school fees and charges have become more important, the educational access of poor rural households, even within a village, may have deteriorated, at least in relative terms and possibly in absolute terms.

The contribution of this paper is to show that education and poverty are closely related in numerous ways, and that the interactions among a set of poverty-related and education-related variables are capable of generating a vicious circle of educational deprivation and poverty, and also a virtuous circle of positive interaction between education and income. The web of hypothesised relationships is estimated empirically for rural China, and the nature of the resultant poverty trap is captured. Their complexity and simultaneous determination in an interrelated system, together with the cross-section nature of our data set, makes it difficult to establish causation, as opposed to association, among our variables. Nevertheless, we believe the results to be original not only for China but also more generally for poor countries, and to be sufficiently suggestive as to provide a road map for further research into the causal mechanisms.

2. Background

China is unusual among poor developing countries in that it is approaching universal basic education. The Law of Nine-Year Compulsory Education, passed in 1986, envisaged that all children would receive this minimum education, commonly 6 years of primary school and 3 years of ‘middle’ (or ‘lower secondary’) school. For implementation, the country was divided into three areas. Cities and economically advanced areas (accounting for 25 percent of the population) were to realise the 9-year target by 1990, areas of middle

development (50 percent of the population) by 1995, but no deadline was set for the economically backward areas (25 percent of the population).

The implementation of the policy is decentralised. In the rural areas county governments have the responsibility to enforce the policy but township (corresponding to the former commune) governments actually carry it out. With economic reform, the fiscal system moved towards decentralisation and self-finance. The majority of government funding for rural education comes from the lowest levels (village, township and county). As fiscal pressures intensified with the decline in redistributive budgetary transfers, rural local authorities resorted to imposing a variety of user fees and charges.

By law primary school is supposed to be free but in practice various fees have been levied: fees for tuition, for textbooks and for uniforms, and 'financial contributions' and other school-based fees or charges. On the basis of ability to pay, we expect to find lower user charges in poorer areas. On the basis of community revenue, however, we expect poorer areas to rely more heavily on user charges. There is likely to be great spatial variation, not only in the quantity but also in the quality of the provision of compulsory education. School quality has emerged as an important policy concern in rural China.

Nearly all children attend their local primary school, normally located in their village, and it is common to attend a middle school outside the village but inside the township (normally comprising several villages). The quality of schooling varies greatly. At one end of the scale there is the 'teaching point' to be found in poor, small, remote localities. This is a one-room school house (often dilapidated) with a single teacher (often with no more than middle school education) instructing all children within walking distance. At the other extreme, schools in the wealthy rural areas may have well qualified teachers and be equipped with libraries, science laboratories, etc.

A study (Liu et al, 2005) based on interviews with 800 primary- and middle-school principals in poor rural areas of Hebei Province in 2002/3 provides evidence that the quality of education is a serious problem. The average number of pupils per class was 49 in primary schools and 61 in middle schools, and in the latter 16 percent of classes had more than 75 pupils. The majority (53 percent) of teachers in primary schools had no more than 12 years of education and training, and 28 percent of middle school teachers were not qualified. No less than 45 percent of teachers in these schools were not being paid the full amount of their salary or were not being paid on time. The problem stems from the decentralised nature of the funding of compulsory education, which was also evident from their survey. Only 7 percent of funds came from central or provincial governments, 11 percent from county governments, 9 percent from township governments, and the remaining 73 percent from village revenues or tuition fees.

At the post-compulsory level - the three years of high (or 'senior secondary') school – there tends to be educational rationing in rural areas. Pupils performing well in examinations can continue to high school, but only if they can afford to pay the school fees. The quality of high school education is in turn important for success in

the competitive national examination that gives access to tertiary education and thence to the high income and urban life that college graduation normally offers.

Taken together, these institutional features suggest that poverty at both the household and the community levels can have an adverse effect on the quantity and also the quality of education that a child in rural China receives.

The data set that we use is the rural component of the national household survey for China, relating to 2002, that was designed by an international team including the authors and organised and administered by the Institute of Economics, Chinese Academy of Social Sciences (IE, CASS). In addition to the main 'Rural Household Questionnaire', there was an 'Administrative Village Questionnaire' and a 'Social Network, Village Affairs and Living Quality' Questionnaire. Both of these latter questionnaires contain questions that were introduced especially for our current purposes.

3. Literature review

Several research papers have explicitly examined the determinants or the consequences of school of enrolment in rural China. We summarise the most relevant briefly for the results that may help in the formulation of hypotheses given our particular research questions and data set. First, we examine the effects of poverty on education.

3.1 The determinants of enrolment

Knight and Li (1996) used the rural sub-sample of a 1988 national household survey to conduct a logit analysis of enrolment for the 14-19 age-group. They found significant coefficients on: male sex (positive), minority status (negative), (predicted) household income (positive, and large in effect), and province mean household income (positive, but small in effect). They posed the question: is rural education demand- or supply-constrained? A second logit, confined to the poorest quarter of households, introduced two new, and significant, variables: opportunity cost of education (negative) and rate of return to education (positive), both measured at the county level. This suggested that the educational decisions of poor households are predictable demand responses to economic incentives. In estimating the determinants of school fees, they found that province income per capita has a positive effect, suggesting that fees represent quality of education or that richer provinces offer higher rates of return to education, for instance because larger subsidies offer better opportunities for progress up the competitive educational ladder. The authors' household income functions also showed the opportunity cost of education to rise and become important between the ages of 14 and 19.

Connelly and Zheng (2003) used a sample of microdata files from the 1990 population census to examine the determinants of enrolment of 10-18 year olds. Their data set is able to control for village fixed effects but lacks many of the plausible explanatory variables, including information on household income and

any variables representing school quality. Their concentration is on primary school: their dependent variables are whether children attended primary school, or graduated from primary school. The variables that increased the probability of attendance or completion include: whether parents attended primary school, whether parents attended middle school, the village school attendance rate, and income per capita in the county (administratively gathered). The presence of siblings decreased girls' chances of schooling.

Brown and Park (2002) specifically examined the effect of poverty on the educational enrolment and outcomes of children aged 5-16, using a 1997 survey of households and schools from poor counties in six provinces. Their measure of household wealth was expenditure per capita (excluding expenditure on education), and they defined a household to be 'poor and credit-constrained' if it is in the bottom third of both expenditure per capita and access to credit. Using a proportional hazard model, they found that children are more likely to drop out of school if the household is poor and credit-constrained (their most important result), if they have fewer siblings, if the father has less education, if the mother had not taken the enrolment decision, and if school fees (possibly proxying school quality)are lower. The sibling result was interpreted as indicating that siblings are complementary rather than competitors for resources, and the fees result as an indication that higher fees improve the quality of education and hence the rate of return. The authors found the test score (for enrolled pupils) to be higher if expenditure per capita is higher (implying that it improves quality), if there are older siblings, and for girls (suggesting that the less able girls drop out of school). However, their variables representing school quality (the pupil-teacher ratio, the proportion of rain-proof classrooms, and proportion of teachers with post-secondary education) had no significant effects on test scores.

Appleton et al. (2006) use the rural sub-sample of a national household survey for 1995 to examine gender differences in educational enrolment rates and educational expenditures. Differences appear only beyond compulsory school age, and so the authors concentrate on the age group 15-18. They find gender differences in enrolment to be particularly pronounced in poorer households. The coefficient on household income per capita is positive and significant for girls but not for boys. This is interpreted as suggesting that girls' schooling is a luxury good whereas boys' is an investment good. Maternal education has a positive and significant effect on enrolment and on educational spending, whereas the effect of paternal education is weaker. The estimation of household income functions provided an economic explanation for the preferential treatment of boys: the coefficients on years of schooling indicate that the return to female education is not different from zero whereas the return to male education, although low, is significantly positive.

De Brauw and Giles (2006) examine the effect of migrant opportunities on high school enrolment in rural China. They use a Ministry of Agriculture panel of households in 52 villages to predict high school enrolment with a set of household and village variables including the number of migrants from the village. This proves not to have a significant coefficient. However, it is plausible that migrant numbers are endogenous to enrolment, e.g. better roads and transport can increase both migration and enrolment, or an adverse income

shock can increase migration but tighten credit constraints on enrolment. The panel provides an instrument for migrant numbers (years since the issue of national identity cards in the village) and the instrumented variable then yields a robust significant negative coefficient. The implication is that migrant opportunities, by raising the opportunity cost of schooling, deter high school enrolment. However, the study lacks good data, such as household income, which might indicate the effects of poverty on schooling.

Gustafsson and Li (2004) used the rural sub-samples of the 1988 and 1995 national household surveys to measure the effect of educational expenditure on poverty. Mean real expenditure on education rose sharply across the deciles of household income per capita in both 1988 and 1995, but especially in 1995, whereas household expenditure on education expressed as a proportion of income fell sharply across the deciles. Real expenditure on education increased rapidly over that period, and in 1995 it represented 3.6 percent of income for the sample as a whole but 7.3 percent for the poorest decile. Using the US$1 PPP poverty line, the headcount index of poverty in rural China fell by 2.2 percent points over the seven years, but with educational expenditures deducted from income, the fall was merely 0.6 percent points. Indeed, the exclusion of educational expenditures made an even greater difference in the officially designated 'poor counties'.

3.2 The effects of education

Secondly, we consider reverse causation in the relationship between education and poverty: the effects of education on income. Yang (2004) examined the contribution of education to rural incomes during the period of factor market liberalisation 1986-1995, when the relaxation of controls permitted households to reallocate resources from agricultural to non-agricultural activities. Using panel data for Sichuan Province, he found that households with a better educated member responded to the new opportunities by devoting more labour and capital to non-agricultural activities that yielded higher returns, and that this contributed to the growth of their household incomes. Thus, the effect of education was indirect: it had no direct effect in the household profit function.

A panel of rural households and workers for the years 1988-1996 was used to argue that education is becoming increasingly important as a determinant of opportunities and income (Zhang et al, 2002). At the first stage of the analysis, a logit analysis estimated the determinants of off-farm work. In 1988 the coefficients of the education variables were insignificant but the village dummy coefficients were important. By 1996 the education terms had become positive, significantly so, and substantial whereas villages had no reliable effect. At the second stage, the education terms were not significant as determinants of non-farm income in 1988, but they had become positive and significant by 1996.

The way in which some Chinese villages get caught in poverty traps and some manage to escape is documented by Knight and Li (1997), who examined several villages in two counties of Hebei province. Villages which were close together geographically could nevertheless be far apart economically. The explanation provided was in terms of factor immobility and processes of cumulative causation. Although a good natural

resource base helped to initiate the process, the main cause of differential village development was non-farm sources of income: migration and village industry. Both were constrained and the easing of constraints involved path-dependent cumulative processes. For instance, migration required a village network of information and contacts, and village industrialisation depended on the accumulation of local skills through a process of learning-by-doing and on the reinvestment of profits. The need for self-reliance meant that village expenditures on infrastructure and education depended heavily on the existence of non-farm activities in the village, but education could itself be important for generating non-farm activities.

Knight and Song (2003, 2005, ch. 8) investigated Chinese peasant choices among three activities: farming, local non-farming and migration. Their source was a rural labour force survey conducted in 8 provinces in 1994. The picture they obtained was one of peasants having a powerful economic incentive to diversify, or move entirely, out of agriculture but being constrained in various ways from doing so. The marginal return to non-farm labour – whether in local or migrant activities – are much higher than the very low marginal return on the farm. Moreover, workers who specialise in non-farm activities work many more days a year than specialist farmers, so enabling them to be more fully employed. Education sharply raises the probability of access to both non-farm activities, monotonically so in the case of village industry. These results suggest that education raises the incomes of rural people by improving their access to non-farm activities, in which conditional incomes are generally higher and in which the conditional returns to education are also higher.

4. Theory and hypotheses: the vicious circle of poverty

4.1 Poverty traps

These research results suggest that we adopt the following theoretical framework and explore the following set of hypotheses using the rural sub-sample of the 2002 national household survey. Our basic hypothesis is that it is easy for households to be trapped in a vicious circle of poverty. Poverty itself can prevent them from grasping the economic opportunities and making the investments which could break the vicious circle and lift them out of poverty. A ‘poverty trap’ or ‘low-level equilibrium trap’ can be a general phenomenon, associated with lack of savings and of access to credit, absence of productive social networks, scarcity of local economic opportunities, and the debilitating effects of a ‘culture of poverty’. For instance, Bowles et al. (2006) argue that there are many conditions that can trap individuals or groups in intractable poverty. These include the existence of critical thresholds, dysfunctional institutions, and neighbourhood effects, each of which can give rise to multiple equilibria. Similarly, Ray (2004), in discussing new research in development economics, concentrates on the potential for there to be more than one equilibrium, and for ‘bad’ equilibria – being socially inefficient although privately profitable - to perpetuate inequality and poverty: the task is then to find policies capable of tipping from one equilibrium to another.

In this paper, however, we concentrate on the potential role of education – its absence or presence - in making or breaking the vicious circle of poverty. Postulating a vicious circle requires justification. There is some to be found in the literature. For instance, Galor and Zeira (1993) theorise that, given initial high inequality, liquidity constraints and set-up costs, it is possible that poor households will not invest in education, so locking their descendants into a poverty trap and perpetuating the inequality. Barham et al. (1995) develop a model in which households differ in their ability to accumulate human capital owing to their facing different rates of return to investment in education and having different capabilities to fund educational expenditures. They show that liquidity constraints can give rise to a poverty trap, in which able children remain uneducated and therefore poor. Durlauf (2006) argues that social interactions can explain persistent inequality across localities. For instance, an absence of role models in a poor locality causes children to develop lower labour market aspirations and thus lower educational aspirations.

The following provides a theoretical framework. We distinguish effects that operate through the individual or the household, on the one hand, and the community, on the other. At the individual or household level, educational enrolment can be constrained by lack of savings and by credit constraints. Household income, which would otherwise be irrelevant to a decision to invest in education, might therefore influence enrolment. The need for additional household current income might require that children work and thus raise the opportunity cost of enrolment. Poverty might reduce the prospective rate of return to education. There are several channels through which this can operate: ill-health, and physical and mental stunting, of the child; weak and unproductive social network to provide economic opportunities; limited productive assets with which to combine human capital; credit constraints which raise the cost of, and thus lower the return to, education; and risk aversion which requires a premium on the return to the investment. For various reasons, the education of parents might encourage the enrolment of their children: their own schooling experience might give the parents a ‘taste’ for education which then motivates their children; and they might impart out-of-school cognitive skills to their children, helping them to perform well at school. If uneducated parents are less successful in these matters, educational deprivation can be transmitted from one generation to another.

At the community level: the poverty of the community might mean that there is no local school, so raising the transport costs of school attendance or making it prohibitive. Inadequate community funding for education can reduce the quality of school provision, so decreasing its prospective rate of return. The rationing of places at post-compulsory levels of education implies that poor school quality at the compulsory levels might prevent access to the high rates of return further up the educational ladder, so deterring even compulsory school enrolment. The lack of local tax revenue can also, or alternatively, necessitate higher school fees being charged, so reducing the prospective rate of return or causing credit constraints to bind. There might also be a low rate of return to education generally in a poor community on account of a lack of local productive economic opportunities. A more general phenomenon, which envelopes these categories, is the ‘culture of poverty’ that can

take hold in a poor community and which dulls supply initiatives and demand incentives. For instance, a low level of education in the community – whether among adults or children - can set a social norm for education and deter individual investment beyond the social norm.

We start with an elementary illustration of a basic poverty trap involving just education, income and a credit constraint. Consider a simple model in which education in years (E) is a function of income (Y) and income in turn is a function of education. What are the conditions required for the existence both of low-level equilibrium poverty trap and also of a high-level equilibrium? Figure 1 illustrates. Assume that the education function $E = E\,(Y,\ldots)$ is inverse-S shaped to reflect the insensitivity of education to income when income is low (the effect of binding wealth and credit constraints), its sensitivity over a range of income, and its dependence only on the rate of return to education when wealth and credit constraints become irrelevant. The income function $Y = Y\,(E,\ldots)$ can be linear, convex or concave without necessarily altering the analysis. In the figure we show it to be concave: income rises with education but at a diminishing rate.

The equilibrium of E and Y at point a is locally stable: small divergences result in a move back to a. However, there is a second equilibrium at point b, which is locally unstable. Thus, an exogenous increase in education from E_1 to a level less than E_2 leads in time, through interaction between the functions, to a restoration of equilibrium: the ensuing rise in income is insufficient to sustain that level of education, and in due course both income and education decline to point a. However, if education can be raised above the threshold, from E_1 to a level beyond E_2, interaction then raises both education and income. A third equilibrium is reached at point c, corresponding to education E_3, which is stable on account of the declining sensitivity and eventual insensitivity of education to income.

The other determinants of Y and of E may be such that some households face only one equilibrium. Thus, a sufficiently large vertical rise in $Y = Y\,(E,\ldots)$ – e.g. through investment in physical capital or an increase in product price - would confine a household to the high-level equilibrium, and a sufficiently large fall would confine it to the low-level equilibrium. Similarly, a sufficiently large vertical fall or rise in $E = E\,(Y,\ldots)$ – e.g. through improved access to credit or an increased subsidy of education, or their reverse, would have the same respective consequences.

This very simple model is intended to illustrate two things. First, it is possible for a household to be stuck in a low-level poverty trap, involving both low education and low income. Second, there is a potential escape from poverty, at least for some households. A large enough intervention to raise education might go through the threshold and propel the household to another equilibrium involving higher levels of education and income. However, in reality, the forces creating vicious and virtuous circles are more complex than that and involve many more interdependent variables.

4.2 The interdependent system

We set up a system of interdependent equations. They are not simultaneous equations in the sense that they are not all contemporaneously determined, involving as they do some relationships from one generation to another. They illustrate the nature of the vicious and virtuous circles that that generate a poverty trap or spring an escape from it. Where

EP	=	parental education
EC	=	community education
YC	=	community income
Y	=	income
EN	=	educational enrolment in years
EQ	=	educational quality
HE	=	health
HA	=	happiness

We have five equations, each dependent on some or all of the others.

Y	=	*Y* (*EN, EQ, HE*)	(1)
EN	=	*EN* (*EP , EC , EQ, Y, YC , HE*)	(2)
EQ	=	*EQ* (*Y, YC*)	(3)
HE	=	*HE* (*Y, EN*)	(4)
HA	=	*HA* (*EN, Y, HE)*	(5)

Figure 2 depicts the same set of relationships, labelling each one with an alphabetical letter. Consider first the direct and indirect determinants of education. At the household level, and within the same generation, there is a positive effect of adult health on household income (**A**), the positive effect of household income on adult health (**B**), and the positive effect of child health on enrolment (**C**). From one generation to the next, household income improves child health (again **B**), child enrolment (**D**) and also educational quality (**E**), and parental education encourages child enrolment (**F**). At the community level, within a generation, community income influences community enrolment (**G**), the quality of education being provided boosts the demand for enrolment (**I**), and enrolment at the community level may encourage enrolment at the individual level (**J**). Across generations, adult educational attainment in the community encourages child enrolment (again **J**), community income influences enrolment (**K**) and improves educational quality (**L**).

There are also outcomes of education, each operating on the generation being educated: educational enrolment of the child raises the income of the subsequent adult (**M**), educational quality also subsequently raises income (**N**), and educational enrolment subsequently improves the health of the person being educated **(O**). Happiness, normally referred to in the literature as subjective well-being or satisfaction with life, depends positively on education (**P**), income (**Q**), and health (**R**).

We intend in this paper to estimate as many of these relationships as possible. Our basic hypotheses are, *first*, that poverty at both the household and community levels adversely affects both the quantity and quality of

education that children receive, and *second*, that the quantity and quality of education help to raise households out of poverty. We intend to explore the numerous education-related processes and mechanisms which can create a poverty trap and which can enable households to emerge from poverty.

This set of hypotheses is already broad and ambitious. Nevertheless, it leaves out several mechanisms by which education can make or break the vicious circle of poverty. The social benefits of education can extend beyond the private economic returns. For instance, education can generate production externalities: the productivity of farm households can be raised by the education of other farmers in the locality (for instance, Weir and Knight 2004, Weir and Knight 2006, for Ethiopia). However, such additional hypotheses fall beyond the scope of this research.

5. Descriptive information

Primary school normally takes 6 years to complete, middle school 3 years, and high school 3 years. No less than 84 percent of children in the sample were enrolled by the age of 7. We therefore take the age group 7-12 to be representative of primary school enrolment, 13-15 to correspond to middle school enrolment, and 16-18 to be relevant for high school. Table 1 shows the enrolment rates in these age groups, for boys, girls and total. We see that the primary school enrolment rate is remarkably high (at 95 percent), as is the middle school enrolment rate (90 percent). All but a small minority of children attend the compulsory years of school. Only in the non-compulsory high school years does the enrolment rate fall sharply (55 percent).

Nevertheless, Table 2 establishes that household income is important to schooling decisions. It uses the detailed information collected on 7-16 year-olds to contrast three categories of 15 and 16 year-olds: those who did not complete their compulsory education, i.e., dropouts before the end of middle school; those who had completed compulsory education but were no longer in school; and those who had completed middle school and were currently in school (refer to the three groups as early dropouts, middle school graduates, and high school pupils). The table shows the five quintiles of households ranked by income per capita. If teenagers are enrolled in school, the household is liable to forgo income and thus its income may be lower. Because school enrolment can reduce household income, we simulate household income to correct for this effect, as will be explained in Section 6. Contrast the poorest households (the first quintile) with the richest (the fifth). Among the children in poor households, 48 percent were early dropouts, 23 percent middle school graduates, and 29 percent high school pupils. The corresponding proportions for those in the richest households were 14 percent, 13 percent, and 74 percent. The ratios of early dropouts to high school pupils are very different: 1.65 and 0.18 respectively. The fourth row of Table 2 shows the enrolment rate of 15-18 year-olds (whatever their level of education). The rate rises monotonically, from 52 percent in the first quintile to 73 percent in the fifth quintile.

In Table 3 we investigate whether household expenditure on each child in school varies with income. In the upper half of the table our criterion for classifying households into the bottom, middle and top thirds is

household income per capita, and in the lower half our criterion is county income per capita, aggregated over all sampled households in the county. We see that both tuition fees and other educational spending (including that on uniforms and textbooks) vary greatly across the three household income groups, and that this is true of all three age groups. In the oldest age group (16-18) the ratio of the top to the bottom income group exceeds 2 to 1. It is the same for both tuition fees and other educational expenditure.

A further point to note from the table is the existence of large differences in expenditure per pupil according to age group. The ratio of expenditures in the age groups corresponding to primary, middle and high school are roughly 1 : 2 : 4. For instance, considering the poorest third of households, the ratios are 1 : 1.81 : 3.64 for tuition fees and 1 : 1.93 : 3.49 for total expenditures. The direct costs of schooling increase substantially as children enter middle school and again as they enter high school.

We expect average county income per capita to serve as a proxy for public expenditure per pupil. If that is the case, the amount that households need to spend might be higher in poorer counties. However, we see from Table 3 that, for household expenditure per pupil, the ratio of the richest to the poorest third of counties generally exceeds 2 to 1. This gives us a strong hint - for further analysis below - that poor households, and also households in poor counties, suffer from poor quality education.

Despite the large differences in household expenditure per pupil, the burden of educational expenditure on household budgets is actually greater for poorer households. The final row of Table 2 expresses total educational expenditure of the household as a percentage of (simulated) household income. We see that this falls almost monotonically, from no less than 26 percent in the poorest quintile to 7 percent in the richest. For the poorest quintile to have to pay a quarter of their income on education is indeed a burden.

6. The determinants of enrolment

Given that the enrolment rates are so high in the age groups relevant to compulsory education, the estimation of enrolment functions predicting enrolment at that level is not very informative. Instead, we show the activity status of children currently out of school (Table 4). The few children of primary school age who are not enrolled are mostly still waiting to attend school (33 percent) or not economically active (at least 40 percent). Among dropouts of middle school age, most (at least 59 percent) are not economically active, but a minority (24 percent) are working. No fewer than 72 percent of the dropouts of high school age are at work, but at least 15 percent are not economically active. Thus, by no means all dropouts find work, at least immediately, but at high school age the choice is largely between employment and enrolment.

The general questionnaire of the rural survey gathered information on the education of all individuals, but the specially designed education module sought more detailed information on the education of children in the age group 7-16. Our analysis combines the two sources. Examining the children aged 15 or 16, there are three groups of interest: dropouts before the end of middle school, dropouts at the end of middle school, and those

continuing after middle school. 25 percent of those who had attended middle school had left school without completing, and of those children (here we can extend the age group to 15-18) who had completed middle school, 41 percent continued to be enrolled in high school. We pose two key questions. *First*, why do some children drop out of middle school before they complete their compulsory education? *Second*, why do some children continue into high school after they complete their compulsory education?

A key variable in our various estimations is income. However, income may well be endogenous, either because it is partly determined jointly with enrolment (omitted variables) or because it is itself partly determined by enrolment (reverse causation). In order to measure the causal effects of income, it is necessary to simulate this variable or to instrument it. We use a measure of income which standardises for the loss of income that enrolment, with consequent withdrawal of labour, might cause: we calculate both farm and non-farm average hourly labour income at the county level, multiply these by the working hours of children of the relevant age group in the household, and subtract this sum from actual household income to get simulated income. This eliminates the difference in household income that is attributable to the greater hours worked by dropouts. Without such a correction, there is a danger that the coefficient of income in an enrolment function will be biased downwards. That is the most likely source of bias. However, it is also possible that parents work harder to fund the education of their children, in which case there can be upward bias. There is a case for instrumenting income in a way that corrects for both sorts of bias. The difficulty, however, lies in finding valid instruments, as nearly all the variables at our disposal might influence the opportunity cost of enrolment. After experimentation we chose to use the identifying variables reported in the notes to Table 6. Generally, however, we use simulated rather than instrumented income in our estimations.

6.1 Explaining middle school drop-out

The implementation of the compulsory education programme is dependent on the financial capacity of local governments. It is inevitable that the programme is more thwarted in the poorer areas. Even in the less poor areas, the rapidly rising tuition fees of recent years may prevent some children from completing compulsory education. Table 5 presents the results of a binary logit analysis estimating the determinants of premature dropping out from middle school. The dependent variable is equal to one if the child had not completed middle school and equal to zero for completion. For each independent variable it shows the coefficient and the marginal effect. Two specifications are shown, one with simulated ln income and the other with five simulated income quintile dummy variables, intended to explore non-linearities in the effects of income.

There are different ways in which to categorise the explanatory variables. One classification is according to demand-side, supply-side, and combined demand-and supply-side factors. However, given our interest in the role of poverty, we classify the determinants as poverty-related at the household level, poverty-related at the community level, other household variables, and other community variables.

Consider first the poverty-related household variables. Simulated ln household income per capita has a highly significant negative coefficient: income reduces the probability of dropping out early (column 1). The marginal effect implies that a fall in income by one standard deviation increases the probability of dropping out early by 4.3 percentage points. As the actual dropout rate is 21.2 percent, this represents an increase of 20.3 percent. However, in the alternative specification (column 2) we see that the relationship is non-linear: only the lowest income quintile has a coefficient significantly different from zero. The large positive coefficient implies that being in the poorest quintile of households (instead of in the middle quintile) increases the probability of dropping out early by 9.2 percentage points, or by 43.4 percent.

Whether the household perceives itself to be subject to a credit constraint is a second indicator of poverty. The coefficient on the variable indicating that the respondent's household would not be able to borrow 5,000 yuan if it was urgently needed is positive but not quite significant. As lack of credit is often a concomitant of low income, we consider also their combined effect: an F-test (for column 1) shows that they are jointly significant at the 2 percent level. An indicator of temporary poverty is provided by the household reporting a natural disaster in 2002. Entered as a dummy variable, it has a significantly positive coefficient: a recent spell of poverty increases the probability of premature dropping out by 5.5 percentage points (25.9 percent). Poverty is indeed an obstacle to the completion of compulsory education.

Parental education is an indirect indicator of household poverty. The coefficients on years of schooling of both father and mother are negative, as expected, but the coefficient is significant only in the case of the mother. For instance, if the mother completed middle school instead of primary school, the probability of early dropout would be diminished by 3.6 percentage points (17.0 percent) (column 1). Thus, the intergenerational transmission of educational outcomes may also perpetuate the poverty of a household. There can be sibling competition for educational resources within a resource-constrained household. We therefore include the possession of a brother and of a sister as explanatory variables, but in neither specification are the coefficients significantly different from zero.

We have four other individual or household variables. The coefficient on the dummy variable indicating that the child is in poor health is positive, as expected, and implies that poor health increases the probability of dropout substantially but the coefficient is not statistically significant, either because there are only 24 cases or because unhealthy children do not get as far even as middle school. We find unexpectedly that being male actually increases the chances of early withdrawal: by 6.1 percentage points (28.8 percent) (column 2). Self-assessment of performance at middle school is highly significant: those who reported that their performance was good had a probability of early dropout 18.7 percentage points (88.2 percent) lower (column 2). We examine below the determinants of middle school performance and find that it is weakened by poverty (Table 9). Finally, we introduce a dummy variable indicating that one or both parents are currently migrants. On the one hand,

migrant parents might be a source of school fees but,on the other hand, schooling might be neglected if parents are absent. However, the coefficient has a different sign in the two specifications and is not at all significant.

Consider how community-level variables affect early dropout from compulsory education. The main unit for analysis is the county (average population about half a million) as only ten households are sampled within each village – too few to provide reliable means. The county (simulated) mean per capita income of sampled households has a positive but insignificant sign. County prosperity has two contradictory effects. On the one hand, local government is better able to subsidise compulsory education; on the other hand, local prosperity raises the opportunity cost of education for teenagers. If it is not a matter of chance, the positive coefficient implies that the latter effect prevails.

Another possible indicator of the opportunity cost of middle school attendance is the migration density in the county: the proportion of workers who are working outside the county is likely to reduce the search and psychological costs of migration. The existence of migration opportunities can have two contradictory effects: it can encourage teenagers to leave school early in order to earn wages as migrants that are higher than they can get in the village, or it can encourage them to stay on at school because the returns to education in migrant employment may be greater than the returns to education in the village. The coefficient is significantly negative: a rise by one standard deviation in the proportion of migrants lowers the probability of early dropout by 3.2 percentage points (15.0 percent). This implies that the latter effect is the more important. As this variable might be endogenous, we use a valid instrument – farm land per capita – to predict migration intensity. However, the coefficient is barely altered, and the Wald test of exogeneity indicates that instrumenting is unnecessary.The distance between the household and the nearest middle school has a positive but slight and insignificant effect, possibly because this distance is typically small (mean 3.8 kilometres).

Certain counties are designated as ‘poor counties’ by central or province governments. Designated poor counties receive special funds from the national or provincial government to help them overcome their poverty. We introduced a dummy variable denoting that a county was designated in that way, to find out whether this support helps to reduce dropout from middle school. In each case the coefficient is negative, as expected, but far from being significant; nor does its inclusion alter the coefficients on income. Designated poor counties are no different in the matter of early dropout from compulsory education. Fiscal support from the higher tiers of government does not appear to help them to achieve universal compulsory education.

The equation contains two variables to proxy the quality of education in the county. One is the proportion of adults aged 25-34 in the county who had attended high school. The coefficient is negative and highly significant: a reduction of the proportion by one standard deviation raises the probability of early drop-out by 5.7 percentage points (26.7 percent) (column 2). This rise may reflect the lower chances of progressing to high school when entry is more restricted or a lower quality of education being received. On average 45 percent

of people reported dissatisfaction with the quality of secondary education in their county. The coefficient on the proportion expressing dissatisfaction in the county is positive, as expected, but not quite significant..

The theoretical concept for assessing the demand for education is the private rate of return to education. We estimated the return to a year of schooling at the community level. It was necessary to estimate the returns separately for farming (normally a household self-employment activity) and non-farming (often an individual wage employment activity), but this was helpful as the estimated returns were distinctly different, with the returns being generally higher in non-farm (province mean 3.4 percent) than in farm activities (county mean 1.2 percent). We estimated both returns at both the county and the province level. In the probit equations the variable for the returns in farming performed better at the county level and the variable for the returns in non-farming performed better at the province level. Thus, it appears that the relevant returns in farming are perceived to be those at the county level, whereas the relevant returns in non-farm activities relate to a larger area, the province, probably reflecting the greater mobility of non-farm labour. The rate of return to education may be important for those deciding whether to complete middle school. The coefficients on the return to a year of schooling in county farming and in province non-farming are both negative, as expected, but neither is statistically significant.

Finally, we note the large and significantly negative coefficient on the county enrolment rate among 15-16 year-olds. A reduction in this enrolment rate by one standard deviation increases the probability of withdrawal from middle school by 6.2 percentage points, or by 29.2 percent (column 2). The interpretation of this effect on individual enrolment is not straightforward (Manski 1993; Soeteven 2006). The problem is to identify the true relationship in the face of the ‘reflection’ problem – whether the mirror image causes a person’s behaviour or reflects it. The county enrolment rate may have an ‘endogenous’ social effect, i.e. a high enrolment of the reference group itself encourages individual enrolment by way of social norms or peer influences or bandwagon effects. The reverse of the coin is that in counties with a ‘culture of poverty’ the accepted social norm for education is low. The difficulty is to distinguish this effect from an ‘exogenous’ effect and a ‘correlated’ effect, both of which can also influence individual enrolment. An exogenous effect depends not directly on the enrolment but on the characteristics of the group, which may be such as to encourage individual enrolment. A correlated effect arises if individuals enrol similarly because they have similar characteristics and face similar incentives to enrol. It is possible that the positive coefficient on the county enrolment rate is due to county characteristics that encourage the enrolment of individuals in the county and which are not included in the enrolment equation., e.g. because the county has a history of strong demand for, and hence supply of, education.

6.2 Explaining high school continuation

Our second exercise is to examine the determinants of the decision to enrol in high school after completion of middle school. For the binary logit analysis, among those 15-18 year-olds who have completed compulsory education, the group still enrolled is given a value of one and the group now out of school a value of zero. Table 6 reports the results, again with two specifications of the income variable. With minor exceptions the same set of explanatory variables is employed. Some of the results are very similar (allowing for the change in sign implicit in the dependent variable) but a few are different.

Consider the variables representing poverty at the household level. Simulated ln income per capita again has a significantly positive coefficient: an increase by one standard deviation increases the probability of enrolment by 18.1 percentage points (column 1); recalling that the mean enrolment proportion is 45.8 percent, this represents an increase of 39.6 percent. The alternative use of (simulated) income quintiles again shows that children in the poorest quintile have significantly worse access to high school, although in this case the probability of enrolment increases monotonically across the quintiles (column 2). Thus, a child in the lowest quintile has a probability of being enrolled in high school 23.2 percentage points (50.8 percent) below, and one in the richest quintile a probability 17.7 percentage points (38.6 percent) above, a child in the middle quintile. Moreover, if a household reports that it cannot raise a loan, there is a significant negative effect: the probability of high school enrolment is reduced by 21.1 percentage points, or by 46.1 percent (column 1). Thus, the poor and credit-constrained are at a huge disadvantage in gaining access to high school.

The education of both parents has a positive and significant effect. Three extra years of schooling (corresponding to a rise in educational level from primary school to middle school, or from middle school to high school) for the father raises the probability of high school enrolment by 4.2 percentage points (9.2 percent), and for the mother by 3.0 percentage points (6.6 percent) (column 1). There is evidence that siblings compete for household resources: the coefficient on the presence of a brother is negative and significant, implying that the probability of continuing to high school is reduced by 7.2 percentage points (15.7 percent) (column1).

Insofar as poverty affects performance at middle school, there is a further mechanism by which the poor are disadvantaged. A report of good performance has a positive and significant coefficient: the probability of continuing beyond middle school is raised by no less than 32.5 percentage points (71.0 percent) (column 1), reflecting the competitive nature of access to high school. As in the case of middle school dropout, the coefficient on poor health has the expected sign but it is not significantly different from zero. However, a more stringent definition of ill-health, creating only 15 cases, does produce a significant and very substantial effect.

There are some interesting contrasts between Tables 5 and 6. In Table 5, against expectations, a boy was shown to be more likely than a girl to drop out of middle school. In Table 6 the significant positive coefficient indicates the opposite: a boy has a 5.6 percentage point (12.2 percent) higher probability of continuing to high school (column 1). Whereas having at least one migrant (away from the county) parent was not significant in the

middle school equation, that coefficient is significant and positive in the high school equation, increasing the probability of enrolment by 12.2 percentage points (26.6 percent). Migrant remittances can be a source of funds for the payment of high school fees and charges; indeed, the need to make such payments might provide a motive for migration.

We turn to the community influences on high school enrolment decisions. County ln income per capita has a significant negative coefficient: an increase by one standard deviation reduces the probability of high school enrolment by 13.3 percentage points (29.1 percent) (column 1). The migration density of the county has a significant positive sign: a rise by one standard deviation raises that probability by 4.3 percentage points (9.4 percent) (column 1). We again instrument migration density using farm land per capita in the county. In this case the Wald test of exogeneity is failed. For instance, there might be some unobserved variable, such as the ambition of local people, that explains both the dependent and the independent variable alike and masks a negative causal relationship. However, in both equations the effect of instrumenting is to raise the positive coefficient (e.g. from the reported 1.413*** in column 1 to 2.522***). Whereas county income appears to proxy the opportunity cost of high school enrolment, the income prospect from migration opportunities does not perform that role.

Considering the variables that might represent the quality of education, both the proportion of adults aged 25-34 who completed high school and the proportion of children at middle and high school who attend 'key schools' (schools that are given more public resources, charge higher fees, and have competitive entry) have positive coefficients, as expected, but neither coefficient is significant. However, the coefficient on the mean level of high school fees in the county is positive, significant and important: a one standard deviation rise increases the probability of high school enrolment by 9.2 percentage points (20.1 percent) (column 2). We take this variable to represent the quality of high school education in the county, and its coefficient to represent the effect of quality on demand. Finally, the coefficient on the proportion of households in the county expressing discontent with the quality of secondary schools is significantly positive. This surprising effect – the opposite of the equivalent result in the equation for predicting middle school drop-out - might reflect the high expectations of those whose children continue to high school and the likelihood that many respondents answer by reference to the local middle school whereas their children attend a more distant high school.

The coefficient on the county rate of return to a year of education in farming is significantly positive. It implies that the probability of high school enrolment rises by 3.8 percentage points (8.3 percent) if the return rises by one standard deviation (column 2). The province rate of return in non-farm activities is also positive but not quite significant. These results suggest that the private demand for high school education is responsive to the prospective returns. Finally, the coefficient on the county enrolment rate for 15-18 year-olds is significantly positive: a one standard deviation increase in that rate boosts an individual's probability of high school enrolment by 16.5 percentage points (35.9 percent) (column 2). This is consistent with a bandwagon effect that

can make or break an educational vicious circle, or with some unobserved factor influencing both dependent and independent variables, e.g. a county tradition of valuing education highly.

We made an attempt to estimate the drop-out and continuation equations using instrumented income instead of simulated income. Unfortunately, estimation with necessarily weak instruments reduced the precision of the estimates and, in the full equation, the coefficients on income were no longer significant. Our solution was to estimate a restricted model, retaining only the explanatory variables that were less likely to be correlated with poverty so that the income variable would be the sole proxy for poverty. The results are shown in Table 6. Poverty does indeed deter completion of middle school significantly: a fall in income by one standard deviation raises the probability of drop-out by 13.5 percentage points (57 percent) (column 1). Poverty also impedes continuation to high school: the same fall in income reduces the chances of continuing by 9.7 percentage points (21 percent) (column 2). In the estimations to come we generally use simulated rather than instrumented income.

7. Further links in the circle of deprivation

The foregoing analysis of the determinants of education suggests that there are further links in the vicious circle of poverty that deserve to be explored. Some of the variables that appeared to be important as determinants of enrolment in Section 6 do themselves need to be explained, so as to discover the underlying rather than the proximate determinants. In particular, we wish to explore whether poverty is a determinant of certain variables that do not themselves represent poverty. These include the county enrolment rate, performance in middle school, health status, the community rate of return to education, and the quality of education.

7.1 Explaining the influence of community

Given the importance of the county enrolment rate in encouraging individual enrolment, we explore the determinants of this variable, using information on the 122 counties in the sample. Table 7 has as the dependent variable the county enrolment rate of those aged 16-18, and as the independent variables a set of county characteristics. The county enrolment rate is positively influenced by the (simulated) income of the county. An income rise by one standard deviation increases the enrolment rate by 5.2 percentage points, i.e. by 9.2 percent. This result is consistent with households being credit-constrained and local governments being budget-constrained. The coefficients on the terms representing the return to a year of education in farming and in non-farming at the county and province level respectively are both positive, as expected, but not quite significant. An indication that the quality of county education is an inducement to enrolment is given by the positive and significant coefficient on the proportion of young adults who attended key schools: a rise by one standard deviation increases the enrolment rate by 3.4 percentage points (6.0 percent). The argument that education begets more education is consistent with the positive and significant coefficient on the proportion of young adults who completed high school: a one standard deviation rise increases the enrolment rate by 4.7 percentage points (8.3 percent). The proportion of workers engaged in non-farm activities in the county appears to raise the enrolment

rate but this effect could arise by chance. In summary, we conclude that the availability and quality of education, and the means to pay for it, are factors which underlie the county enrolment rate, which in turn plays an important role in determining household enrolment decisions.

The theoretical criterion for the enrolment decision is the private rate of return to education. Non-farm activities are normally carried out on an individual basis whereas farm activities are normally carried out by households. Because this distinction requires different methods of estimation (as will be explained in Section 8), the returns are reported separately for farm and for non-farm activities. In any case, the annual returns at the county level are very different, averaging 1.4 percent in farming and 4.0 percent in non-farming. Table 8 reports the determinants of the county-level returns to a year of schooling. We see that mean household income per capita in the county has a significantly positive effect on the non-farm rate of return, possibly reflecting the greater productivity of human capital in a more developed local economy. Thus, it appears that the deterrent effect of mean county income per capita in the enrolment equations (Tables 4 and 5) reflects a high opportunity cost and not a low benefit of education. Taking the county mean value, household educational expenditure per capita has a significantly positive effect on the rate of return to education in farming. This suggests that the quality of education can be important in determining this rate of return, and it indicates a mechanism by which low quality of education can deter enrolment. Thus, the poor may be doubly disadvantaged. However, we cannot rule out reverse causation: a high rate of return may induce spending to improve educational quality; nor can we rule out the possibility that counties in which households spend little on education also suffer from a culture of poverty which deters complementary out-of-school human capital acquisition. Another indicator of educational quality, the proportion of adults in the county who attended key schools, is not a significant determinant of either rate of return, possibly because the proportions attending key schools and also remaining in the county are too small.

7.2 Explaining school performance

Respondents were asked about their school performance relative to their peers when at middle school or, if they had not attended middle school, their performance when at primary school. We analyse the replies of children aged 13-18. Half of them (50 percent) reported that their performance was either good or very good. This group takes the value one, and the rest the value zero, in our binary logit analysis. Such a self-reported indicator may be inaccurate and also biased (few respondents owned up to a bad or very bad performance) but our estimates produce plausible and statistically significant results (Table 9). We see that gender has no significant effect on performance, but both father's and mother's years of schooling have a significant positive effect. For instance, if the father completed middle school instead of primary school, the probability of performing well rises by 5.1 percentage points, or by 10.2 percent. When ln simulated household income per capita is included as a proxy for parental support for the child's education (column1), it has a positive and

significant coefficient: a one standard deviation rise in this variable increase the chances of performing well by 2.3 percentage points, or by 4.6 percent. However, when the highly correlated variable, ln educational expenditure, is included as well, the income coefficient is not significant and it is the educational expenditure variable that does the work (column 2). In other words, income has its main effect through enabling parents to spend more on their child. The marginal effect for ln expenditure indicates that an increase by one standard deviation increases the probability of performing well by 5.6 percentage points, or by 11.2 percent. The two variables with links to household poverty have significantly negative coefficients (column 1). A dummy variable indicating that the child works after school and another indicating that the child is not healthy reduce the probability of good performance by 4.1 and 19.7 percentage points respectively.

7.3 Explaining the determinants of educational quality

One important link is the role of educational quality. The enrolment functions contain explanatory variables that are serving as proxies for quality of education: these suggest that higher quality raises enrolment. However, the proxies are in turn influenced by the poverty of the household or the locality. The implication is that quality and quantity of education may interact positively, and that households can get stuck in a poverty trap because they experience both poorer quality and lesser quantity of education: the value of their human capital suffers in two ways.

The quality of education that children receive can depend on the intensity with which they study. Moreover, children may need to study hard at primary and middle school if they are to succeed in the competition for entry to high school. If children from poor households are more likely to engage in work out of school, they may have less time for study and therefore may learn less and also have a smaller chance of securing a rationed high school place. This issue is explored in Table 10. We know whether certain activities are 'always performed after school': farmwork, childcare, and housework. The table reports binomial logit estimates for participation in these activities. We concentrate on the variables which are both relevant to the hypothesis and statistically significant. We see that (simulated) income significantly reduces the probability of both farmwork and housework. There is also a negative effect when at least one parent is a migrant, suggesting that parental migration has its effect more through remittances than by leaving children to cope with household tasks. The variables indicating the quality of the child's school – school fee, and attendance at a key school – each significantly reduces both farmwork and housework. The same is true of the variables that proxy the general quality of education in the county – village educational expenditure aggregated to the county level, and the proportion of county students aged 15-18 attending key schools. The table also shows that at least two of the activities are more likely if the child is older, or a girl, or comes from a minority group, or has younger siblings. In summary, Table 10 suggests that the value of human capital received by poor children is adversely affected by out-of-school work in two ways. There is the direct effect of household poverty in reducing the amount of study

time available, and the indirect effect which operates, at both household and county level, through the quality of schooling being received. Children lack the incentive to study hard if they cannot learn much at school, and if the prospect of continuing to high school is low.

We have information on whether there is a primary school in the village and on the nature of that primary school. We had expected to find that poor children are less likely than others to have a village primary school and more likely to have a single classroom school rather than a full grade primary school. However, the poorest quintile of households, and the poorest quintile of villages, are not at a disadvantage in these respects. In fact, the proportion of households in villages without a primary school is actually lower for the poorest quintile of villages (11 percent) than for the richest quintile (30 percent). However, this result is likely to reflect the latter's willingness to travel for higher quality education, probably in larger, better resourced, schools.

Even in rural areas, children can face a choice of schools of varying quality. Schools may charge different fees and involve variation in other costs, including transport costs, uniform costs, and costs of materials and books. Independently of the school, parents can choose to boost their child's education by buying more books and paying for extra tuition and other extracurricular activities. The quality of education as opposed to mere school attendance may be important both for the value of human capital acquired in school and for the prospects of entering the generally competitive and rationed system beyond compulsory education. The amount that parents spend on education may thus have a powerful effect on future outcomes.

One of the issues underlying the quality of education is whether private and public expenditures on education are substitutes or complements. For given enrolments, if educational quality is constant, more public funding implies less private funding: there is less need to charge school fees. By contrast, if quality is a variable to be optimised, greater public funding might encourage households to spend privately on education: the required condition is that government spending raises the marginal benefit of private spending. A third possibility is that both government and household spending are similarly influenced by a third factor such as local prosperity. For instance, if poor households are credit-constrained and local governments in poor counties are budget-constrained, poverty can depress educational quality. Higher incomes in a county might then enable both local governments and households to increase their educational spending so as to bring educational quality towards the optimum level, which is governed by the rate of return to quality-enhancing expenditure.

We analyse the determinants of household educational expenditure per child, conditional on the child being enrolled. Table 11 reports the results of OLS equations in which ln educational expenditure for children at primary, middle, and high school are the dependent variables, concentrating on the explanatory variables that illuminate the relationships between poverty at the household and community levels and educational quality. The (simulated) income variable has significant positive coefficients at all three levels. Parental migration, plausibly reflecting disposable income in the form of cash remittances, also significantly increases the amount spent, but significantly so only at middle school. The older sibling variables reduce expenditure significantly at high school

level, and the younger sibling variables do so at all three levels. This evidence of competition for household resources means that households are unable to spend as much as they would like on education. If households were not constrained, expenditure should reflect the rate of return to the investment and should not be influenced by the presence of siblings. A child's attendance at a key (secondary) school involves substantial and significant additional expenditure. There is a strong and significant positive coefficient on county income per capita and on county educational expenditure per capita, each coefficient rising monotonically with educational level. The former is a proxy for the capacity of local governments to pay for education and the latter a direct but crude measure of actual payment. These results indicate that household expenditure is complementary to local government expenditure on education, and not a substitute for it. The implication is that, if schools are of higher quality, parents are in turn willing to invest more in educational quality for their children.

We can learn more about the nature of this complementarity, and of its relationship to poverty, by distinguishing between household expenditure on school fees and on other educational inputs: expenditure on books, on uniforms, and a residual input. The most important of these components is the residual, which is likely to include out-of-school tuition. Table 12 reproduces Table 11 but distinguishes between fee and non-fee expenditures and converts the dependent variables into logarithmic form so as to assist comparison of relative sensitivity. Note that total educational expenditure rises from 426 yuan per child at primary school, to 882 yuan at middle school, and to 2,566 yuan at high school, and that school fees represent 48, 46, and 49 percent of the total respectively. An increase in simulated income by one standard deviation raises total educational expenditure per enrolled pupil by 2.7, 3.7, and 4.3 percent at primary, middle, and high school level respectively. The household income variable is no less important for other expenditure than for tuition fees. Other expenditure appears to be more sensitive than tuition fees to an increase in county income per capita at the primary and middle school levels but less so at the high school level. Local government educational expenditure per capita has more influence on household spending on tuition fees than on other items in middle schools.

In summary of Tables 11 and 12, household educational expenditure per child rises with income per capita of the household, income per capita of the county, and also the county average of local government educational expenditure per capita. This is generally true of both tuition fees and of other, more discretionary, educational expenses. It appears that household and community prosperity improve both the quality of education provided in the schools and the quality of additional educational support provided by the parents, and that greater subsidisation by local government is associated with more, rather than less, household expenditure on a child's education.

8. The effects of education

A potentially important link in the circle of deprivation is whether, and to what extent, the outcomes of education are less favourable for the poor than for the less poor. This effect can operate both through the poorer

quality of education that they receive and the poorer economic opportunities that they face. The main objective of this section is to examine the economic effects of education, concentrating *first* on whether, and by what mechanisms, there are positive returns to education in rural China and, *second*, on whether household or community poverty reduces the prospective returns to investment in education.

8.1 Education and allocation of labour between farm and non-farm activities

The descriptive Table 13 shows how workers are allocated between farming and non-farming activities and, within the latter, between local and migrant activities. We classify farm workers and non-farm workers according to which activity they spend more time in, and local non-farm and migrant workers according to whether they are working in the county. Our purpose is to contrast poor households, and young workers, with others.

In the sample as a whole, we see that the farm/non-farm percentage division is 64/36, and the local/migration division of the 36 percent is 22/14. However, there is a marked difference between the poorest and the richest third of households: 75 percent of the former are in farming but only 52 percent of the latter. There is little difference in the importance of migration but a large difference in the importance of local non-farm work: 12 percent and 34 percent respectively. A similar pattern is found in comparing the poorest and the richest third of counties, except that there is a larger difference in migration (favouring poorer counties), probably reflecting the role of migrant networks in fostering migration. We see again the heavy dependence on farming (73 percent) of the poorest half of households in the poorest half of counties.

The relevance of age is seen in the row showing labour allocation among workers aged under 30 years. Non-farm activities are much more important for the young (accounting for 49 percent of workers) than for the old, especially migrant labour (accounting for 30 percent). For workers under 30 in the poorest half of households in the poorest half of counties, local non-farm opportunities are rare but migration is common (28 percent being migrants).

A further indication that the distinction between farm and non-farm activities is important for income is provided by the association between household income per capita decile and the distribution of activities. The proportion of non-farm hours worked to total hours worked (equal to 37 percent overall) rises monotonically from 14 percent in the lowest decile to 65 percent in the highest (table not shown). Table 14 illustrates the reasons why and how this powerful association arises. It shows that the ratio of the average return per hour of labour in non-farm to that in farm activities is 1.91. At the margin, the ratio is far higher, at 11.92. We also see that the ratio of the return per hour of labour in local non-farming to that in migration is 1.51 on average and 1.37 at the margin. These disparities are consistent with there being labour market segmentation and rationing of non-farm jobs, as argued by Knight and Song (2005, ch.8). If so, does education assist households to earn non-farm income?

Educational decisions are liable to be based on the prospects facing young people. They face a roughly equal division of labour between farm and non-farm activities, and even those in poverty have reasonable prospects of non-farm employment. The role of education in determining whether people engage in farming or non-farming, and the returns to education in farming and non-farming, are therefore important issues. We explore them in turn, the hypothesis being that education can have both allocative and efficiency income benefits.

Table 15 reports the relevant results of a multinomial logit analysis to predict the determinants of activity choice, the reference activity being farming. Men are more likely than women to be found in both local and migratory non-farm activity. The propensity to migrate falls with age but the chances of being in a local non-farm job initially rises with age. Being healthy is important for both activities but more so for migration. The proportion of workers in the county who are migrants and the proportion who work locally off-farm have predictable effects: the migration of others greatly increases the chances of own-migration, probably through network effects, and high non-farm labour density is a sign of local employment opportunities.

Education raises the chances of being in either of the non-farm activities, but especially if it is local rather than outside the county. For instance, the marginals imply that high school rather than primary school completion increases the probability of migrant work by 33 percentage points and of local non-farm work by no less than 64 percentage points. Thus, education is a lifeline for young people who wish to escape from the farm, whether their motive is economic or social. Very similar general results are obtained from equivalent analyses conducted on the sub-samples of workers in the poorest third of households and workers aged under 30.

8.2 The economic benefits of education in farm and non-farm activities

The benefits of education arise from two effects: an allocation affect and an efficiency effect. Our estimation strategy is as follows. It is necessary to instrument hours worked on account of its likely endogeneity: more hours may be worked if the returns are higher. We first estimate farm and non-farm working hours separately. A Tobit estimation is necessary owing to non-participation in an activity: the proportion of censored observations are 15 and 54 percent respectively for farm and non-farm work. The instrumented hours variable can then be included in the second-stage income functions which measure the benefits of education in farm and non-farm activities.

Consider first the non-farm equation, based on data for individuals currently working but under 65 years of age. The key explanatory variables are the education categories, chosen rather than years of schooling so as to permit non-linearities and to assist policy analysis. We regress working hours also on self-rated health status, gender, age, age squared, and the county average hourly wage in farm and non-farm work. Table 16 presents the results: the coefficients, the marginals, and the decomposition of the marginals into that part due to change in the

number of hours worked given positive non-farm hours and that part due to change in the probability of working off-farm. All the coefficients except age are highly significant.

We concentrate on the explanatory variables that are relevant to our hypothesis. The mean value of non-farm hours worked is 732 hours per annum (14.1 hours a week). With no education as the omitted category, the education dummy coefficients are all positive and rise monotonically with education level. For instance, having high school instead of primary school education raises hours worked by 345 hours, two-thirds of this rise being due to the increased chance of participation. Education appears to encourage or enable workers both to participate in non-farm activities and also to work longer hours in them. A report of being in good health or very good health raises non-farm work by 109 hours. As expected, the county average non-farm wage raises hours worked and the corresponding farm wage lowers them, hours decline throughout the working life, and men work longer hours than women. Two demographic variables, the number of children aged under 16 and the number of old people over 65 in the household, which should not influence household income directly, are included as instrumental variables so that instrumented hours can be included in the income equation to come.

An equivalent estimation for farm hours worked by individuals is reported in Table 17. Again, the education dummies and the health variable are the most directly relevant. The equation differs only in that three household variables are added: the amount of farm land cultivated, productive fixed assets, and the total number of labourers on the farm. The mean value of the dependent variable is 956 hours per annum (18.4 hours a week). All the coefficients except productive fixed assets (a poor proxy for specifically farm equipment) are highly significant. With no education again as the omitted category, the education dummy variables are all negative and rise monotonically in negative value. For instance, having high school instead of primary school education reduces hours worked on the farm by 110, two-thirds of which is due to a lower probability of working on the farm. This is the most important result of Tables 16 and 17: as education increases, so workers switch strongly to non-farm work and weakly from farm work. This result indicates the allocative benefit of education.

Reporting good health increases farm work by 30 hours. As expected, a higher county average non-farm wage, by raising the household's supply price, reduces farm hours. We also include the predicted number of non-farm hours worked, as we expect non-farm work to be preferred if it can be obtained, and farm hours then to be adjusted. The coefficient is indeed significantly negative but its impact is small, the marginal having a value well under unity, suggesting that, with many rural households suffering from underemployment, non-farm work can be expanded without contracting farm work equivalently or even substantially. Farm hours continue to increase with age until the mid-forties.

Table 18 reports the determinants of the non-farm income of individuals. Our dependent variables are absolute non-farm income, to assist the simulations to come, and also ln non-farm income. Consider column 2: income is higher for men than for women, and rises with age until the late forties. Being healthy raises income by 12 percent. The coefficients on the education dummy variables, beyond primary school, are large and highly

significant, and rise monotonically with education level. In the absolute income equation (column 3), the differences in coefficients show the marginal product of each education level. For instance, the marginal product of middle school is 481 yuan and that of high school is 417 yuan per annum. However, this understates the value of education to a worker because education also increases the number of non-farm hours worked. When the hours variable is excluded from the specification, the marginal product of middle school becomes 835 yuan and of high school 781 yuan per annum (equation not shown).

The first column of Table 18 introduces three proxies for the quality of education that workers had received, each measured at the county level: the proportion of young adults who completed high school, the mean level of school fees per pupil, and the proportion of households that expressed dissatisfaction with the quality of secondary education. Each variable has the expected sign, and each is statistically significant when introduced singly. When they are entered together, they are jointly highly significant but, owing to their collinearity, only school fees remain significant. An increase in the county mean fee by one standard deviation raises non-farm income by 11.5 percent; and if all three variables are increased by one standard deviation, the gain is 22.0 percent. This suggests that the quality of education improves productivity in non-farm activities.

Because farming is a household activity and a household contains workers of different educational levels, a different specification of the income function is required. In addition to farmland, productive fixed assets, and the proportion of household workers who are healthy, we include the number of (instrumented) farm hours worked by household workers at each educational level. The differences in the coefficients on these education-specific hours variables thus provide the marginal product per 1000 hours of each education level.

Table 19 reports the results. As expected, farm land and productive fixed assets raise farm income. Poor health does not have an adverse effect, probably because there is plenty of underemployed household labour to draw on. All the education variables of interest are positive, highly significant, and rise with education level. The exception is college education: very few college graduates remain in the rural areas, and those who do remain are unlikely to farm except as a hobby. When we introduced the three county variables representing the quality of education that workers had received, as in Table 18, none proved to be significant (equation not shown).

The marginal product, for 1,000 hours of farm work, of middle school is 44 yuan and of high school 26 yuan (column 2). To make these figures comparable with those for non-farm activities, we multiply up in the proportion 1,519 (the average number of non-farm hours worked) to 1,000: the marginal products become 67 and 40 yuan respectively Clearly, the marginal product of education in farming is much lower than in non-farming. There are efficiency benefits of education in both farm and non-farm activities but they are considerably greater in the latter.

8.3 Simulating the effects on poverty of raising educational levels

It is possible to make use of the estimates in Tables 17-19 to conduct a simulation analysis. The counterfactual question being posed is: what are the effects on the incidence of income poverty of improving the educational level of the rural labour force? We recognise that this is a simplistic exercise. We abstract form the long time lag between educational expansion and the consequent improvement in the income generation process, and from all the other relationships involving education and poverty for which this paper educes evidence. We also ignore the possibility that the intervention will alter the estimated relationships, e.g. by increasing competition for non-farm jobs among the educated, and the possibility that the economy will change before those being educated enter the labour market, e.g. by increasing the availability of non-farm jobs.

We use two poverty lines, corresponding to the $1 a day and $2 a day concepts. Converted into the 2002 prices in rural China, these become 925 yuan and 1,850 yuan. Table 20 shows the results of this exercise. The first column reproduces the actual situation. It reports actual average farm income, non-farm income, other income and overall income. The mean numbers of farm and non-farm hours are also shown. Finally, it shows the actual proportion of rural households for which income per capita is under the $1 a day line (10.9 percent), and the proportion under the $2 a day line (43.0 percent).

The first simulation is to assume that those with no education and those with primary education had completed middle school, i.e. compulsory education. The combination of the allocative and efficiency effects is to reduce 'headcount' poverty to 2.6 percent ($1 a day) and 27.3 percent ($2 a day).Our other simulation is to assume that every worker with education below high school instead had completed high school. The poverty rates fall dramatically, to 0.5 percent ($1 a day) and to 9.0 percent ($2 a day). These benefits stem from a reallocation of labour towards non-farm activities and the higher returns to education in such activities.

8.4 Explaining differential returns to education

The estimated returns to education in rural, or farming, activities in the developing world tend to be positive but low in absolute terms and in relation to the returns in urban, or non-farming, activities. Phillips (1994) in a meta-analysis of the returns to education in farming in developing countries - covering 30 studies and 59 data sets - estimated an average rate of return to an additional four years of schooling of 9.5 percent, implying an annual return of 2.4 percent. In China, Li and Zhang (1998) found the returns to a year of education for farmers in Sichuan Province in 1990 to be 3.3 percent (average household education) or 2.7 percent (highest household education).We shall show below our own estimates of the return in farming to be variously 1.5 percent (all rural households, however little they farm), 3.2 percent (traditional farming households) and 4.8 percent (modern farming households) a year. The returns to education for farmers appear to be relatively low, both in China and more generally in poor countries.

The returns to education can be lower for poorer households either because the quality of their education is inferior or because their opportunities or resources are inferior. These adverse effects can operate at the household level or at the community level. Poor households may be unable to afford high quality education and may be limited in their economic opportunities by lack of resources, of information, and of ambition. In poor counties the quality of educational provision may be low, local economic opportunities may be scarce, and the local resource base or infrastructure may be weak. We hypothesise that poor households, and households in poor counties, face lower returns to their education.

Table 21 shows the estimated returns to a year of schooling in farming, non-farming and combined activities, distinguishing between different sub-samples according to household or county income per capita. It is not clear what education variable should be used in the case of household production because knowledge is in principle transferable and available within the household. After experimentation we decided to use the greatest years of education among the farm workers (for farm income) and among all workers (for combined income) in the household. Conditioning variables are kept to a minimum because of the possibility that education works partly through them.

In the sample as a whole the return to a year of education is much lower in farming (1.5 percent) than in non-farming (10.8 percent), and for household income as a whole (6.4 percent). The overall return corresponds to a weighted average of the two component returns, with the weight on non-farming rising with educational level (as was shown in Tables 15 and 16). Including province dummy variables (and so estimating the within-province returns), the pattern is similar although in each case the return is a bit lower; we report only the equations which do not condition on province.

Our interest is in the contrast between the average, the poor and the rich. The estimates of the returns to schooling for income-based sub-samples are subject to downward bias owing to truncation, but comparisons of the returns for the different income groups are valid. In all three cases, the return for the poorest third of households is lower than that for the richest third, being less than half in the farming and combined cases. For instance, the overall return to a year of education is a paltry 2.0 percent for the poorest third of households and 4.5 percent for the richest third. Poor households thus have less incentive than the non-poor to invest in education. Households in the bottom third of counties face distinctly lower returns in farming than do those in the top third. There are differences in returns both for poor households and for households in poor counties, suggesting that the problem must be addressed at both household and county levels.

We search for the underlying reasons for the differential returns to education. The returns to education are likely to depend positively on the level of technology being used in the production process. Accordingly, they are likely to be greater not only in rural non-farming than in farming but also, within the latter, greater in ‘modernising’ farming conditions than in a ‘traditional’ farming environment. Argument in support of this hypothesis is to be found in e.g. Schultz (1975), and summary evidence in a meta-analysis which found an

average return per year of schooling of 2.9 percent in modern farming and of 1.9 percent in traditional farming (Phillips, 1994, p.155).

We test the latter hypothesis in the following way. First, we identify 'farming' households. Our criterion is that the household derives the majority of its income from farm activities. Second, we distinguish between modern and traditional farming households on the basis of their relative emphasis on the traditional farm activity, grain production. Our necessarily crude definitions are that households which obtain the majority of their farm income from grain production are 'traditional' and those which obtain a minority are 'modern'. We then estimate a farm income function, with three key explanatory variables: average years of education of the household workers, a 'modern' dummy variable, and a modern* years of education interaction term. The dependent variable is ln household farm income and the conditioning independent variables are land area used, hours of labour on the farm, and productive fixed assets.

Table 22 reports the results. All the coefficients are significant. Being a modern farming household adds 33 percent to income, and the rate of return to a year of education is 4.2 percent (column 1). However, when the modern* years of schooling interaction term is introduced (column 2), the coefficient on the interaction term is 2.1 percent per year. The implication is that the return to a year of schooling is 2.7 percent among traditional farmers and 4.8 percent among modern farmers. This is powerful evidence that education is more valuable for farming households which are able and willing to diversify away from traditional crops.

The subsequent analysis has shown that our simulation analysis of the effects of increasing education on poverty – being based on average relationships in the sample as a whole - may be misleading. Insofar as poor households face lower returns to education, the poverty gains will be overstated. Insofar as young people face better prospects of non-farm employment, the poverty gains will be understated. Accordingly, we redo the simulation analysis: first, for the poorest third of households and, secondly, for households of which the head is aged under 30 years. The results are different for two reasons. Not only are the values of the relevant characteristics of each sub-sample different from those of the sample as a whole, but also the equations are different. In both cases we re-estimate the equations in Tables 16-19 - on which the counterfactual simulation analysis is based - using only the sub-sample observations.

Table 23 presents the results for the poorest third of households. The poverty rate is extremely high: 32 percent are below the $1 a day line and 100 percent below the $2 a day line. Assuming that all workers have completed at least compulsory education has a trivial effect on the poverty rate at the higher line but reduces the rate at the lower line by over a third. Giving all workers at least a high school education again has little effect on $2 a day poverty but reduces $1 a day poverty by almost two-thirds. Despite the poverty trap, the poorest can be helped out of their poverty by education.

Table 24 reports the simulations for the sub-sample of young households. The baseline is quite similar to that for the sample as a whole: even though young individuals spend more time off the farm, their households

actually work fewer non-farm hours. However, the simulated falls in their poverty rates are somewhat greater, reflecting the strong tendency for the educated young to find non-farm employment.

8.5 Education, poverty and health

Individuals were asked to classify themselves on a five-rung health ladder, from very healthy to very unhealthy. 84 percent reported being healthy or very healthy: we distinguish this group of adults aged 16-65 (taking a value of zero in the binary logit analysis) from the 16 percent whose health was so-so, bad, or very bad (taking a value of one). The explanatory variables are a dummy for male sex, age and age squared, a series of educational level dummies, and (instrumented) ln income per capita. All the variables predicting poor health have the expected sign and are significant at the 1 percent level (Table 25). An individual is less likely to be unhealthy if he is male or young. The two variables of interest are the individual's education and the household's income. There is a significant fall in the chances of being unhealthy with educational level. For instance, having completed high school instead of middle school reduces that probability by 2.5 percentage points, i.e. by 14.4 percent; and an increase in household income per capita by one standard deviation reduces it by 1.1 percentage points, i.e. by 6.5 percent. Thus, both income-poverty and education-poverty induce ill-health in adults. A binary probit equation to predict ill-health among children aged 0-15 was unable to find links to the poverty or the education of their parents: only male sex and age were significant, both reducing the probability of being unhealthy.

8.6 Education, poverty and innovation

Education may be important in fostering risk-taking behaviour in agriculture (Knight et al, 2003), and poverty may itself foster a culture of poverty in which enterprise, risk-taking, and innovation are repressed. The survey contained a question asking the respondent about their attitude to adopting new agricultural technology. The answers ranged from very positive to not at all positive. We create a dummy variable for 'positive' or 'very positive' replies, the other three replies being the reference category. In a binary logit equation, the explanatory variables are male sex, age and age squared, and years of schooling of the respondent, and ln household income per capita. We consider the sub-sample of 'farming' households, i.e. with more than half of their income from farming. The hypothesis is that both the education and the income variables have positive coefficients.

Table 26 shows this indeed to be the case. When ln household income per capita is used, the coefficient is positive and highly significant (0.251), and raising the income variable by one standard deviation increases it by 3.4 percentage points, or by 4.9 percent (equation not shown). However, causation is questionable. For instance, an inhospitable terrain may produce both poverty and a negative attitude. In the table we therefore instrument the income variable. The coefficient on income is reduced to 0.088 and it is no longer significant. Raising education from primary to high school increases the probability of having a positive attitude to

innovation by 13.1 percentage points, or by 19.6 percent. This result, also, is open to criticism: more able people may be both more educated and more positive towards innovation. Correction of this potential bias would require a valid instrument for years of education. Nevertheless, the results of our attitudinal analysis suggest yet another link in the chain: low education and low income might be associated with the negative attitudes that are part of a 'culture of poverty'.

8.7 Education, poverty and subjective well-being

Our final exercise is to examine the relationship between our central variables education, income, and health and a variable that can be regarded as providing a broader criterion than any of these for assessing the quality of life: happiness or subjective well-being. We are encouraged in this exercise by the rapidly growing literature on the economics of happiness which generally finds powerful regularities – involving statistically significant coefficients with the hypothesised signs – in many data sets. It is arguable that 'subjective well-being poverty' is an encompassing concept into which income poverty and capabilities poverty can be incorporated (Kingdon and Knight, 2006). Our hypotheses are that, in improving subjective well-being, education, income and good health all help to reduce poverty defined in this encompassing sense. Any definition of poverty involves a value judgement on the part of the researcher, and subjective well-being poverty at least has the virtues of being based on individual choice, of concern for subjectively perceived misery, and of measurability.

Survey respondents were asked how happy they were nowadays: very happy, happy, so-so, unhappy, or not at all happy. We convert this information into two forms of dependent variable, a binary variable identifying those reporting themselves to be happy or very happy, and a cardinal variable ranging from very happy = 4 down to not at all happy = 0. The independent variables of most interest in our subjective well-being functions are the respondent's years of schooling, a dummy variable indicating that the respondent reports being unhealthy, and ln income of the respondent's household. The conditioning variables are whether the current living standard is reported to be lower, or higher, than five years ago, age and age squared, and dummies for male sex, marital status, and – to standardise for temporary effects - whether the respondent's current mood is good.

The choice of dependent variable makes no difference to the results, in line with the methodological conclusion of Ferrer-I-Carbonnel and Frijters (2004). In both cases all the relevant coefficients are significant at least at the 5 percent level (Table 27). The coefficient on years of schooling is positive. In the OLS equation, raising education from primary to high school increases the happiness score (mean value 2.68) by 0.036; in the logit equation it increases the probability of being happy or very happy by 3.0 percentage points, or by 4.8 percent. This is of course in addition to the indirect effects of education on happiness via income and health. Insofar as causation runs from (past) education to (current) income and health, the full effect of education can be shown by omitting the latter variables. On their omission from the equation (not shown), the effect of education

is more than doubled: additional education raises the happiness score by 0.114 and the probability of being happy by 7.1 percentage points, or by 11.4 percent.

When ln household income per capita is included, its coefficient has a significant positive value in the OLS equation (not shown) of 0.177. However, because income is potentially endogenous, we instrument the income variable: the OLS coefficient rises to 0.278, again highly significant. It appears that some unobserved variable (such as a driven personality) adds to income but subtracts from happiness. An increase in ln income per capita by one standard deviation increases the happiness score by 0.127 and the probability of being happy by 7.4 percentage points, or by 11.9 percent. Being in poor health decreases happiness by 0.237 points and the probability of being happy by 13.7 percentage points, or by 22.0 percent.

The conditioning variables behave as expected: we see the importance of comparisons with previous living standards, implying that relative income is relevant as well as absolute income; and the importance also of standardising for current mood. Age has the commonly found U-shaped relationship with happiness; men report being less happy than do women, and marriage is good for happiness. The crucial results, however, are that education, income and health each has the hypothesised effect on subjective well-being.

9. Summary

9.1 The seventeen relationships

Our first objective in this section is to present the various results in a systemic way, showing that, on account of their many interactions, the whole is greater than the sum of the parts. Only by considering all the relationships together can a coherent argument be developed. We take equations 1 – 5 and Figure 2 as our theoretical framework. Consider the evidence on each of the hypothesised relationships **A**,…, **R** in turn.

A. Being in good health raises the income of the individual and of the household. Self-reported good health increases the number of hours worked by individuals in non-farm activities by 14.9 percent and in farm activities by 3.2 percent (Tables16 and 17). Standardising for hours worked, good health raises a worker's non-farm income by 11.7 percent but has no significant effect on the household's farm income (Tables 18 and 19).

B. The logit equations predicting the health of an adult show that a one standard deviation fall in ln household income per capita raises the probability of being unhealthy by 1.1 percentage points, or by 6.5 percent (Table 25). However, we lack evidence that parental poverty affects child health adversely.

C. The logit equations predicting enrolment show that ill-health prejudices a child's completion of middle school, increasing the chances of dropping out by 4.3 percentage points, or by 20.3 percent (Table 4), and of

continuation to high school, the reduced chances in this case being 3.4 percentage points, or 7.5 percent (Table 5). However, neither effect is statistically significant, probably because there are few cases and because ill-health deters children from reaching even middle school.

D. We have evidence that household poverty, as measured by income per capita and the existence of a credit constraint, has adverse effects both on completion of middle school and on continuation to high school. The logit equations predicting enrolment show that a one standard deviation decrease in ln household income per capita increases the probability of dropping out from middle school by 4.2 percentage points, or by 20.0 percent (Table 4), and decreases the probability of continuing to high school by 18.1 percentage points, or by 39.5 percent (Table 5). Our other indicator of poverty, suffering a credit constraint, reduces the chances of completing middle school (Table 4). The effect is not statistically significant but, given the close relation between poverty and lack of credit, our finding of their joint significance is relevant. Lack of credit significantly reduces the chances of continuation to high school, by 19.1 percentage points, or by 39.6 percent (Table 5).

E. There is evidence that lower parental income means inferior educational quality for their children. We see this in Tables 11 and 12, where income per capita raises the total household expenditure per enrolled child, and where there is evidence of sibling competition for educational spending. An increase in household income by one standard deviation raises total household educational expenditure per enrolled child by 2.7, 3.7 and 4.3 percent at primary, middle, and high school respectively. Income poverty and sibling competition increases the amount of time that children spend in farmwork and housework after school (Table 10).

F. There is also evidence that the education of children suffers if their parents are poorly educated. The mother's education significantly encourages completion of middle school: if she has middle school instead of primary school education, the chances of dropout are reduced by 3.9 percentage points, or by 17.0 percent (Table 4). The education of each parent encourages continuation to high school: an additional three years of education for the father increases the chances of continuation by 4.2 percentage points, or by 9.2 percent (Table 5).

G. We found evidence that community enrolment is positively related to community income. A rise by one standard deviation in county income per capita raises the county enrolment rate by 5.2 percentage points, or by 9.2 percent (Table 7).

J. Community enrolment has a strong association with individual enrolment. A fall in the county relevant enrolment rate by one standard deviation reduces the chances of completing middle school by 6.4 percentage

points, or by 30.2 percent (Table 4). The same fall reduces the chances of continuing to high school by 16.5 percentage points, or by 36.0 percent (Table 5). Whatever the mechanism by which this effect operates, educational poverty of the individual is closely linked to educational poverty in the community.

K. Community income is a determinant of individual enrolment. In this case, however, the evidence runs counter to the hypothesis of a vicious circle of education- and income-poverty. Higher income in the county has no significant effect on the chances of completing middle school (Table 4). However, a fall by one standard deviation in county income per capita increases the chances of continuation to high school by 13.3 percentage points, or by 29.0 percent, suggesting that the opportunity cost of high school is lower in poorer counties (Table 5).

L. The community also influences the quality of education that children receive. Some of the evidence on this is indirect and inconclusive, being based on differential returns to a year of schooling: a rise in county income per capita significantly raises the returns in non-farm activities and a rise in county educational expenditure per capita significantly raises the returns in farm activities (Table 8). There is a strong effect, at all three school levels, of household income per capita, county income per capita, and the county average of village educational expenditure per capita on the household's expenditure per enrolled child (Tables 11 and 12). Our interpretation is that households demand a higher quality of education, and are prepared to pay for it, as their own income rises and as the quality of publicly provided education improves.

I. The quality of education that a child receives, or expects to receive, has an effect on enrolment. We know that school performance is raised by expenditure on a child's schooling, which in turn is influenced by household income per capita (Table 9), and that school performance is a powerful determinant of enrolment. Reported good performance in middle (or primary) school decreases the probability of drop-out from middle school by 19.0 percentage points (Table 4) and increases the probability of continuing to high school by 32.5 percentage points (Table 5). The proportion of respondents in the county who are dissatisfied with the quality of secondary schools is significant in Table 4: a rise by one standard deviation in the proportion dissatisfied increases drop-out by 5.2 percentage points. Mean high school fees in the county appear to be irrelevant at middle school level but have a positive effect on enrolment at high school level. A rise by one standard deviation in county high school fees – which we take to imply higher quality - raises the probability of high school enrolment by 9.2 percentage points (Table 5).

M. Education is found to raise individual and household income. It does so through an effect on hours worked and on the income per hour, in farm and non-farm activities. The non-farm sector pays much better than does

the farm sector, both on average and at the margin (Table 14). A rise in the education of a worker from primary to high school level increases the number of hours worked in the non-farm sector by 47.1 percent (Table 16), and reduces those worked in farm activities by 11.5 percent (Table17). Standardising for hours worked, the marginal products of high school and of middle school are 417 and 481 yuan respectively in non-farming (Table18), and the corresponding marginal products are 40 and 67 yuan respectively in farming (Table 19). One of the ways in which education can raise income is by inculcating a positive attitude to innovation. Raising education from primary to high school level increases the probability of having a positive attitude to adopting new agricultural technology by 13.1 percentage points, or 19.6 percent (Table 26). As against these benefits of education, the fact that the returns to schooling are lower for low-income households and low-income counties helps to create and maintain a poverty trap (Table 21). Similarly, the returns to schooling are lower for traditional farming households than for those engaged in 'modern' farming activities (Table 22).

N. We have indirect evidence that the quality of education that workers have received is a determinant of its economic benefit and therefore of their income. One indication is the willingness of parents to spend more on the education of their children the higher is their own income per capita and the higher the income per capita, and the average village educational expenditure per capita, in the county (Tables 11 and 12). This suggests that parents perceive a benefit from improved educational quality. Moreover, if all three county variables that we take to proxy the quality of education that workers have received are together increased by one standard deviation, although farm income is unaffected, non-farm income is raised by 22.0 percent (Table 18).

O. Workers' educational levels have an effect on their health status. Having completed high school instead of middle school, or middle school instead of primary school, reduces the probability of ill-health by 2.5 percentage points, or by 14.4 percent (Table 25).

P. The logit equation predicting the determinants of subjective well-being imply that an increase in education from primary school to high school level directly raises the probability of being happy or very happy by 3.0 percentage points, or by 4.8 percent (Table 27). This increase is 7.1 percentage points, equivalent to11.4 percent, if the indirect effects of education, working through its influence on income and health, are included as well.

Q. The same equation shows the effect of income on subjective well-being to be powerful. An increase in ln income per capita of the household by one standard deviation raises the probability of the respondent's being happy by 7.4 percentage points, or by 11.9 percent (Table 27).

R. Finally, from the same source we see that poor health has a serious adverse effect on subjective well-being: it reduces the probability of being happy by 13.7 percentage points, or by 22.0 percent (Table 27). Thus, given that subjective well-being is a criterion for poverty, we see that providing education, increasing income, and improving health can all reduce the risk of being in poverty.

9.2 Can education break the vicious circle of poverty?

We can now see why and how a poverty trap can exist. Low income restricts investment in education by households. Low education of the parents deters enrolment, as does low education in the community. Both low income and low education adversely affect health status; poor health status, in turn, reduces income and deters enrolment. Low income of the household and also of the community reduces the quality of education. Low quality of education in turn deters enrolment. Both low quantity and low quality of education reduce the income benefits of education. Low income, low education, and poor health all reduce subjective well-being, which can be viewed as an encompassing indication of poverty. Some of these relationships are short term but others are transmitted across generations, implying that household poverty can be persistent over the years.

The interaction among the variables which has the potential to create a vicious circle also has the potential to create a virtuous circle, provided that a critical threshold can be surmounted. Using the simple analogy of Figure1, we see the following possibilities. There may be a single equilibrium at the (lower) point *a*, in which case the interactions simply move that equilibrium point upwards. There may instead be two stable equilibria, corresponding to points *a* and *c*, and an unstable equilibrium at point *b*. The issue is whether a shock or intervention is sufficiently large to move the variables from the low-level equilibrium, *a*, beyond the critical threshold, *b*, so that they converge on the high-level equilibrium, *c*. We cannot tell from the evidence which of these cases is relevant. Nevertheless, consider four shocks or interventions.

First, assume an exogenous shock which raises the household's income. This has the potential to encourage enrolment and also to improve the quality of education demanded. Over a generation, these gains feed through into higher household income; this in turn improves health, which itself has knock-on effects on education and income. It also raises the quantity and quality of education of the next generation, with further indirect effects. Secondly, assume an exogenous shock that raises not only the income of the household but also that of its community. In addition to the consequences listed in the first example, there are benefits accruing from the greater public revenue of the community: higher quality of education, the possibility of a demonstration effect from a higher community enrolment rate, and the possibility of a higher rate of return to education as the structure of the local economy changes. By contrast, we also saw that the higher income of the community deters enrolment by raising the opportunity cost of school attendance.

Thirdly, consider a policy intervention from above which is intended to improve the quantity and quality of education and is aimed at poor households. It will take time for the benefit from the educational improvement

to flow, in the form of increased household income, but eventually this rise will in turn improve health, with its beneficial effects, and also improve the quantity and quality of education received by the next generation. Fourthly, consider the same policy intervention, now aimed at poor communities. If all households in the community are targeted, in addition to the benefits accruing to each household described in the third case, there can be demonstration effects of the higher enrolment rate. If local governments are targeted, the intervention does not require the mediation of household demand to improve the quantity and quality of education supplied.

Insofar as educational interventions raise income and ease credit constraints, on the one hand, and improve human capital, on the other, the ensuing increased access to funds for investment and enhanced profitability of investment may create further indirect effects through increased physical capital formation.

10. Conclusions

Section 9 has already provided a systemic summary of the paper. We draw three types of conclusion: first, concerning methodology and further research; second, on the relationships between education and poverty in a poor country, and how they throw light on the persistence of, and potential escape from, a poverty trap; third, on the implications of the research for policy in general and for rural China in particular.

Most research papers in economics test one or two hypotheses: the research and its conclusions are specific and narrow. Broader conclusions might be drawn by introducing these research relationships into a more general system – by placing the results in the context of the research literature on related topics – but there is a natural reluctance to venture out in that way.

In this paper we have taken a different approach. Our hypothesis is a very general one, which in turn gives rise to many sub-hypotheses, each of which requires empirical testing. In arguing the case we have had to adopt a broader, and therefore necessarily shallower, approach than is conventional in the research literature. The paper is also inevitably longer than is usual for journal articles. The trade-off is worth it because light cannot otherwise be thrown on an important general phenomenon viewed as a whole. Our combination of broad hypothesis and empirical estimation has not, to our knowledge, been previously attempted on this topic.

The main lesson for research is the potential importance of educational quality – both for human capital formation and for the labour market benefits of education – and its relationship to poverty. Our findings suggest a need for further research on the causes and consequences of school quality. This in turn points to the use of tests of reasoning ability and cognitive achievement in the estimation of both educational achievement production functions and of human-capital-augmented income functions (an early example of which, for poor countries, was Boissiere et al., 1985).

The main hypothesis of the paper is that there exist a set of relationships between income, or other indicators of poverty, and education. These run in two directions – from income to education and from education to income – and they interact among themselves in many ways. The upshot is that these processes can generate

reinforcing upward or downward pressures; in the latter case a vicious circle involving a low-level equilibrium for both education and income.

No fewer than 17 hypothesised relationships **A**,…, **R** were estimated, and then brought together in Section 9, where they were summarised individually and presented collectively to form a system. Given the inter-relatedness of the many variables in play, and the cross-section nature of the data set, it is difficult to isolate causation, as opposed to association, in some of the relationships. However, that task is less important in the present context than establishing that there is an inter-related and mutually reinforcing system of relationships. This set of relationships constitutes our evidence that an education-poverty, income-poverty trap can exist. Unobserved heterogeneity can pose a problem for establishing causality, but if omitted but correlated variables such as lack of personal 'ability' or 'a culture of poverty' are themselves determinants of poverty, they simply strengthen the poverty trap.

Causation becomes important for understanding the underlying reasons for the problem and for assessing the effects of policy interventions. Knowledge of the causal relationships is crucial in devising policies that will engineer an escape from the poverty trap. For instance, in the absence of a good instrument for schooling, we cannot be sure that education will have the powerful effects on income that are implied by our estimates. For that, further, more detailed, research is required. An underlying policy issue is whether expanding educational enrolment alone would be sufficient or whether this should be accompanied by complementary policy interventions. These might aim to raise the prospective rate of return to education by, for instance, improving the quality of schooling, or improving opportunities in the local economy, or weakening a debilitating culture of poverty.

Given a positive exogenous shock or policy intervention, there are two possibilities. One is that the change may simply raise the position of the low-level equilibrium. The other is that interaction among the variables may set in train a process of cumulative causation and a virtuous circle generating a high-level equilibrium of both education and income. We cannot establish from our evidence that there are two equilibria and that a sufficiently large shock can move the system from one to another. Nevertheless, our findings of numerous positive relationships among a set of interacting variables opens that possibility. In any case, nothing in our general argument or policy implications hinges on the issue of whether there are two equilibria or just one, movable, equilibrium.

The theory of a low-level equilibrium and the possibility of escape from it through a process of cumulative causation were pioneered fifty years ago (e.g. Leibenstein, 1957). Although never dormant, it has become popular again through endogenous growth theory, often involving human capital (e.g. Lucas, 1988, and applied to poverty by Azariadis, 2006) and the new emphasis on understanding persistent poverty (e.g. Ray, 2004; Bowles, 2006). This paper provides some empirical support – rather lacking in the literature – for the role of education in the persistence of, and potential escape from, a poverty trap.

We turn to the policy implications of our findings. The issue for the poor in rural China is not only the quantity but also - and probably more so - the quality of education. Recall that the enrolment rates for 7-12, 13-15, and 16-18 year-olds are 95, 90, and 55 percent respectively. It is true that we found income and credit constraints to be important in determining dropout from middle school and continuation to high school. However, enrolment decisions are partly governed by the quality of schooling that has already been received and the quality that is expected.

There are huge differences in expenditure on the education of a child, both by local governments and by households, and these differences are in turn closely related to community and household income. The household differences occur both in school fees and in other educational expenditure such as books, uniforms, and private tuition. The quality of education, much influenced by expenditure, has its effect not only on the demand side (by influencing the perceived rate of return to education) but also on the supply side (by determining which children succeed in the competition for high school places). Thus, even in countries where the quantity of education causes serious policy concern, the quality of education may deserve no less research- and policy-attention.

Consider the policy implications for rural China. The underlying problem for people in poor households, poor villages, or poor counties is the degree of fiscal decentralisation to be found in rural society. Chinese peasants are effectively expected to 'pull themselves up by their own bootstraps'. The solution to the education-poverty, income-poverty trap requires institutional reform: greater fiscal centralisation and equalisation – a theme that has already been stressed by the authors (Knight and Li, 1999). There should be more redistribution of tax revenue from higher to lower tiers of government, and from lower tiers of government to households. These redistributions should be aimed at the poor.

The evidence suggests that the interventions should be made both at the household-level and at the local community-level. At the household-level, Chinese policy-makers might wish to introduce a version of the so-called Progresa scheme, pioneered in Mexico in 1997 and subsequently adopted in several other Latin American countries. An advantage of such schemes is that they lend themselves to experimental interventions designed to measure their effects accurately: the findings so far have been promising (Skoufias and McClafferty, 2001).

The Progresa scheme is intended to address extreme poverty in rural areas of Mexico by developing the human capital of the poor. The recipients are poor families in poor communities. The scheme encourages school enrolment by providing conditional subsidies: it is necessary for parents to send their children to school. By contrast, in the Chinese conditions of high enrolment rates, such a scheme might well have its main effect on the quality of education that households demand; or if necessary the scheme could be adapted so as to promote the quality of education. The lower tiers of government could similarly receive conditional revenue transfers, the requirements of which would involve taking measures to improve the quality of education that local governments provide.

The Chinese government has indeed begun to move in this direction. As part of the recently declared policy of promoting the 'Harmonious Society', the educational burden on rural households has been progressively lightened. In 2001 the central government began to provide free textbooks for poor pupils attending compulsory education (years 1-9) in officially designated poor counties. In 2005 it introduced a policy of also exempting them from the payment of tuition fees and of providing them with accommodation subsidies. Government funding for exemption from tuition fees was extended to all pupils in western provinces in 2006 and to all provinces in 2007. Thus there are no longer tuition fees for compulsory education in rural China but measures to maintain and improve the quality of education may also be needed.

References

Appleton, S., Knight, J. & L. Song (2006). Why do girls in rural China have lower school enrolment? *World Development*, 34(9),1639-53.

A. Costas (2006). The theory of poverty traps. What have we learned? In Bowles, Samuel et al.(eds), *Poverty Traps.* Princeton: Princeton University Press.

Boissiere, M. , Knight, J. & Sabot, R. (1985). Earnings, schooling, ability and cognitive skills. *American Economic Review*, 75, 1016-30.

Bowles, S. , Durlauf, S. N. & K. Hoff (eds) (2006). *Poverty Traps.* Princeton: Princeton University Press.

Brown, P. & Park, A. (2002). Education and poverty in rural China. *Economics of Education Review*, 21, 523-41.

De Brauw, Giles, A. and J. (2006). Migrant opportunity and the educational attainment of youth in rural China. wordprocessed.

Connelly, R. & Zheng, Zh. (2003). Determinants of school enrolment and completion of 10 to 18 year olds in China. *Economics of Education Review*, 22 (4), 379-88.

Durlauf, S. N. (2006). Groups, social influences and inequality. In Bowles, Samuel et al (eds), *Poverty Traps.* Princeton: Princeton University Press.

Ferrer-I-Carbonnel, Frijters, A. & P. (2004). How important is methodology for estimates of the determinants of happiness? *The Economic Journal*, 114 (497), 640-59.

Gustafsson, B. and Shi, L. (2004). Expenditures on education and health care and poverty in rural China. *China Economic Review*, 15, 292-301.

Kingdon, G. & Knight, J. (2006). Subjective well-being poverty versus income poverty and capabilities poverty? *Journal of Development Studies*, forthcoming.

Knight, J. & Shi, L. (1996). Educational attainment and the rural-urban divide in China. *Oxford Bulletin of Economics and Statistics*, 58(1), 81-115.

Knight, J. & Shi, L. (1997). Cumulative causation and inequality among villages in China. *Oxford Development Studies*, 25(2),149-172.

Knight, J. & Shi, L. (1999). Fiscal decentralization: incentives, redistribution and reform in China. *Oxford Development Studies*, 27(1), 5-32.

Knight, J. and Song, L. (2005). *Towards a Labour Market in China.* Oxford: Oxford University Press.

Knight, J., Weir, Sh. & Woldehanna, T. (2003). The role of education in facilitating risk-taking and innovation in agriculture. *Journal of Development Studies*, 39(6),1-22.

Leibenstein, H. (1957). *Economic Backwardness and Economic Growth.* New York: John Wiley.

Liu, F. (2004). Basic education in China's rural areas: a legal obligation or an individual choice? *International Journal of Educational Development*, 24, 5-21.

Ming, L., Bing, Y. & Damin, W. (2006). Investigation and thinking compulsory education in poor rural areas. *Forum of Contemporary Education* …? (in Chinese).

Tianyou,L. & Zhang, J. (1998). Returns to education under collective and household farming in China. *Journal of Development Economics*, 56, 307-35.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22, 3-42.

Manski, C. P. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, 60(3), 531-42.

Phillips, J. M. (1994). Farmer education and farmer efficiency: a meta-analysis. *Economic Development and Cultural Change*, 43(1), 149-65.

Ray, Debraj (2004). What's new in development economics? In Szenberg, Michael and Lall Ramrattan (eds), *New Frontiers in Economics*. Cambridge: Cambridge University Press.

Schultz, T. W. (1977). The value of ability to deal with disequilibria. *Journal of Economic Literature*, 13(2), 827-96.

Skoufias, E. & B. McClafferty (2001). Is Progresa working? Summary of the results of an evaluation by IFPRI', *FCND Discussion Paper No. 118*, Washington D.C., International Food Policy Institute.

Soeteven, A. (2006). Empirics of the identification of social interactions: an evaluation of the approaches and their results. *Journal of Economic Surveys*, 20(2), 193-228.

Weir, Sh. & Knight, J. (2004). Externality effects of education: dynamics of the adoption and diffusion of an innovation in rural Ethiopia. *Economic Development and Cultural Change*, 53(1), 93-114.

Weir, Sh. & Knight, J. (2006). Production externalities of education: evidence from rural Ethiopia. *Journal of African Economies*, forthcoming.

West, L. & Wong, C. (1995). Fiscal decentralisation and growing regional disparities in rural China: some evidence on the provision of social services. *Oxford Review of Economic Policy*, 11(4), 70-84.

Yang, D. (2004). Education and allocative efficiency: household income growth during rural reforms in China. *Journal of Development Economics*, 74, 137-162.

Zhang, L., J. Huang & Rozelle, S. (2002). Employment, emerging labour markets, and the role of education in rural China. *China Economic Review*, 13, 313-328.

Figure 1. A poverty trap

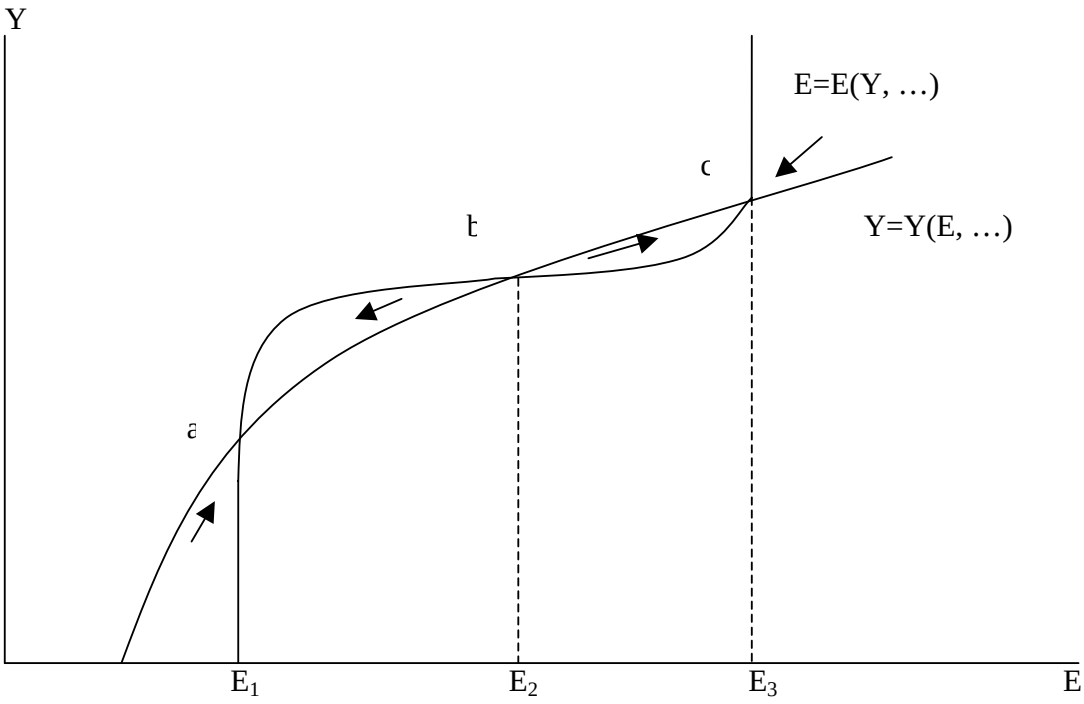

Figure 2. Vicious and virtuous circles

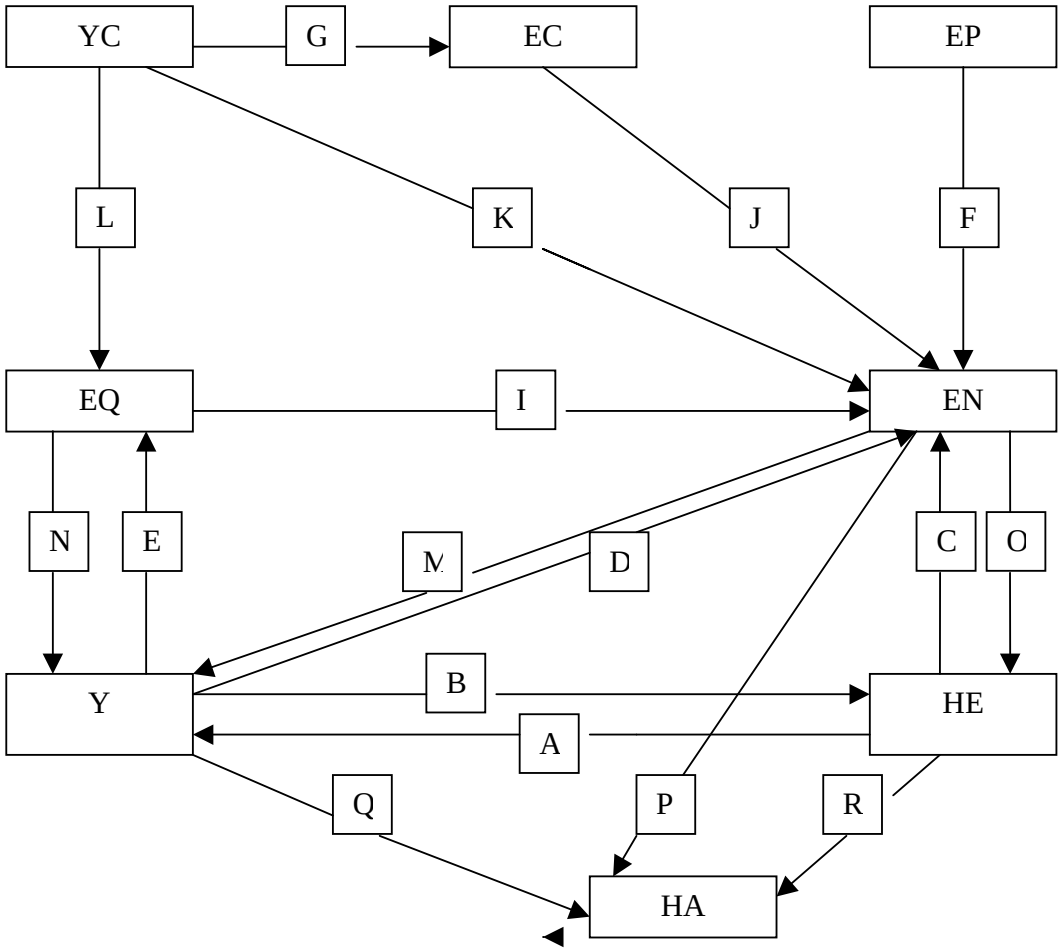

Table 1. Educational enrolments and expenditures by income per capita quintile

	Income quintile					Total
	1	2	3	4	5	
Distribution of 15-16 year-olds by educational status (column percentage)						
early dropouts	47.6	21.4	19.4	19.9	13.5	25.8
middle school graduates	23.5	20.7	23.1	15.1	12.8	19.2
high school pupils	28.8	57.9	57.5	65.1	73.7	55.0
Enrolment rate of 15-18 year-olds (percentage)	52.3	62.0	62.8	65.4	72.5	62.7
Educational expenditure as a proportion of household income (percentage)	25.5	13.9	9.9	8.5	7.1	13.0

Notes:
1. The categories excluded from the distribution of 15-16 year-olds are the negligible few who have never attended school and the children who are still enrolled in middle school: their choices will be made later.
2. Simulated income rather than actual income is used throughout (see section 6).

Table 2. Educational expenditure per child in school, in the bottom, middle and top third of households by household income per capita and by county income per capita, by age group

	Household income per capita			
	bottom third	middle third	top third	ratio of top to bottom
Age group 7-12				
tuition	171.5	234.5	289.8	1.69
other	167.0	207.8	317.8	1.90
total expenditure	338.5	442.3	607.6	1.79
Age group 13-15				
tuition	310.2	405.5	551.1	1.78
other	344.2	410.1	538.9	1.57
total expenditure	654.4	815.4	1090.0	1.67
Age group 16-18				
tuition	623.8	975.1	1478.3	2.37
other	557.5	777.8	1254.6	2.25
total expenditure	1181.3	1752.9	2732.9	2.31

County income per capita

Age group 7-12				
tuition	152.8	232.3	327.8	2.14
other	156.6	197.2	334.5	2.14
total expenditure	309.4	429.5	662.3	2.14
Age group 13-15				
tuition	279.4	429.1	593.4	2.12
other	319.3	393.5	609.4	1.91
total expenditure	598.7	822.6	1202.8	2.01
Age group 16-18				
tuition	733.1	946.0	1466.0	2.00
other	583.2	770.8	1260.8	2.16
total expenditure	1316.3	1716.8	2726.8	2.07

Table 3. Status of children who are currently out of school, by age group, column percentage

	Age group		
	7-12	13-15	16-18
Working	6.3	23.7	72.3
Unemployed	29.1	19.7	10.2
Disabled	4.0	1.8	1.0
Pre-school	32.6	0.4	0.0
Housework	1.1	6.6	4.5
'Dropout'	10.9	39.1	4.5
Other	26.9	8.7	7.5
Observations	175	274	1,258

Table 4.The determinants of drop-out from middle school, binary logit estimates

	1		2	
	coefficient	marginal	coefficient	marginal
simulated ln income per capita	-0.479***	-0.055		
simulated ln income per capita				
quintile 1			0.698 *	0.092
quintile 2			-0.210	-0.023
quintile 3			0.000	0.000
quintile 4			-0.030	-0.003
quintile 5			-0.423	-0.044
credit-constrained	1.031	0.165	1.076	0.174
disaster in 2002	0.458**	0.055	0.451*	0.053
father's years of education	-0.011	-0.001	-0.014	-0.002
mother's years of education	-0.108**	-0.012	-0.114**	-0.013
one or more brothers	0.080	0.009	0.117	0.013
one or more sisters	0.071	0.008	0.085	0.010
male	0.515**	0.058	0.545**	0.061
good performance in middle school	-1.638***	-0.183	-1.695***	-0.187
one or both parents a migrant	0.016	0.002	-0.009	-0.001
county simulated ln mean income	0.611	0.070	0.519	0.059
county migrant density	-2.244**	-0.257	-2.266**	-0.256
distance to nearest middle school	0.026	0.003	0.024	0.003
county proportion of adults 25-34 who completed high school	-5.230***	-0.598	-5.493***	-0.621
county proportion dissatisfied with secondary schools	0.406	0.046	0.356	0.040
county rate of return in farming	-27.862	-3.186	-29.164	-3.297
province rate of return in non-farming	-6.695	-0.765	-4.850	-0.548
county enrolment rate for 15-16 year olds	-3.011***	-0.344	-3.041***	-0.344
designated poor county	-0.210	-0.023	-0.266	-0.028
ill-health	0.287	0.036	0.196	0.024
intercept	1.686		1.649	
mean of dependent variable	0.212		0.212	
pseudo R-squared	0.254		0.256	
number of observations	631		631	

Notes: 1. We use simulated rather than instrumented income; the latter has a similar but slightly smaller coefficient on ln income per capita and has less statistical significance.

2. Standard errors are corrected for clustering at the county level.

Table 5. The determinants of continuation from middle school to high school: binary logit estimates

	1		2	
	coefficient	marginal	coefficient	marginal
simulated ln income per capita	0.775***	0.191		
simulated ln income per capita				
quintile 1			-1.003***	-0.232
quintile 2			-0.068	-0.017
quintile 3			0.000	0.000
quintile 4			0.356**	0.088
quintile 5			0.715***	0.177
credit-constrained	-0.967**	- 0.211	-0.912**	-0.202
father's years of schooling	0.057*	0.014	0.064**	0.016
mother's years of schooling	0.040*	0.010	0.042*	0.010
one or more brothers	-0.293**	-0.072	-0.281*	-0.069
one or more sisters	-0.168	-0.041	-0.157	-0.039
good performance in middle school	1.362***	0.325	1.362***	0.325
male	0.229**	0.056	0.240**	0.059
one or both parents a migrant	0.489***	0.122	0.490***	0.122
county ln simulated mean income	-1.128***	-0.278	-1.028***	-0.254
county migration density	1.408**	0.347	1.333***	0.329
county proportion of adults 25-34 who completed high-school	1.197	0.295	0.988	0.244
county proportion of children in key schools	0.599	0.148	0.460	0.113
county mean school fees	2.166***	0.534	2.257***	0.557
county proportion dissatisfied with secondary schools	0.461***	0.114	0.450***	0.111
county rate of return in farming	15.195**	3.744	15.524**	3.830
province rate of return in non-farming	5.759	1.419	5.446	1.344
county enrolment rate for15-18 year olds	3.840***	0.946	3.799***	0.937
designated poor county	-0.029	-0.007	0.025	0.006
ill-health	-0.141	-0.034	-0.092	-0.023
intercept	-6.312***		-5.887***	
mean of dependent variable	0.458		0.458	
pseudo R-squared	0.250		0.243	
number of observations	1,822		1,822	

Notes: As for Table 4.

Table 6. The determinants of drop-out from middle school and the continuation to high school, restricted model with income instrumented

	drop-out		continuation	
	coefficient	marginal	coefficient	marginal
income (instrumented)	-0.656**	-0.202	0.385**	0.148
good performance in middle school	-0.838***	-0.244	0.811***	0.307
male	0.020	0.006	0.119**	0.046
county rate of return in farming	-8.154	-2.506	13.611***	5.222
intercept	0.231		-1.168***	
Wald chi-sq	85.58		264.97	
number of observations	753		2262	

Notes: 1. The Stata program ivprobit maximum likelihood is used.
2. The identifying variables in the income equation are: whether the village leader has experience of running a business, whether the village had to sell grain at the officially specified price in 2002, and whether the terrain of the village is plain (the reference category), hilly or mountainous. All four variables have the expected sign and are statistically significant.
3. The test of over-identifying instruments (indicating that if one instrument is valid then at least one other instrument is also valid) is passed in the continuation but not in the drop-out equation. In both equations, the Wald test of exogeneity indicates that we cannot reject the hypotheses that income is exogenous

Table 7. The determinants of the county enrolment rate for those aged 16-18, OLS estimates

Simulated county income per capita	0.040**	0.054***
Return to a year of education in farming in the county	1.958	2.154
Return to a year of education in non-farming in the province	1.286	1.552*
Proportion of pupils 16-18 in the county who attend key schools	0.905**	1.136***
Proportion of adults 25-34 in the county who completed high school	0.482**	
Proportion of the labour force in the county who are in local non-farming	0.160	0.191
Constant term	0.236***	0.243***
Adjusted R-squared	0.336	0.276
Mean of dependent variable	0.563	0.563
No of observations	122	122

Note: ***, **, and * denote statistical significance at the one per cent, five per cent and ten per cent levels respectively.

Table 8. The determinants of the county-level returns to education: OLS estimates

	Farming	Non-farming
household income per capita in county	0.000	0.005***
household educational expenditure per capita in county	0.225**	0.132
county proportion of adults aged 25-34 who attended key schools	0.009	-0.004
intercept	0.011***	0.026***
mean value of dependent variable	0.014	0.040
adjusted R-squared	0.037	0.053
number of observations	122	122

Notes: 1. County educational expenditure is village educational expenditure per capita with weighted aggregation to the county level.
2. ***, ** and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.

Table 9. The determinants of school performance of children aged 13-18, binary logit analysis

	Coefficients		Marginal effects	
	(1)	(2)	(1)	(2)
Male	-0.055	-0.048	-0.014	-0.012
Father's years of schooling	0.068***	-0.061***	0.017	0.015
Mother's years of schooling	0.059***	0.048***	0.015	0.012
Ln simulated household income per capita	0.121**	0.046	0.030	0.012
Ln expenditure on child's schooling		0.231***		0.058
Child works after school	-0.165**	-0.088	-0.041	-0.022
Child not in good health	-0.826***	-0.844***	-0.197	-0.200
Constant term	-1.685***	-2.589***		
Pseudo- R-squared	0.028	0.027		
Mean of dependent variable	0.504	0.500		
Number of observations	3,564	3,323		

Notes: 1. The dependent variable is (self-reported) excellent or good performance at middle school (or, if middle school was not attended, at primary school) = 1.
2. The dummy variable for work is child does farmwork or housework after school during school terms = 1, and the dummy variable for health is child's health is so-so, bad, or very bad = 1.
3. ***, ** and * denote statistical significance at the one, five and ten per cent level respectively.

Table 10. Logit model indicating that a child always performs certain activities after school: children aged 6-16

Dependent variable:	Farmwork	Childcare	Housework
Simulated income per capita	-0.005***	0.000	-0.005***
at least one parent a migrant	-0.026*	-0.019**	-0.016
attendance at key school	-0.060***	-0.017	-0.100***
school fee (000 yuan)	-0.029***	-0.003	-0.031***
age (years)	0.024***	-0.004***	0.024***
male	0.010	-0.024***	-0.068***
minority group	0.100***	0.008	0.154***
younger sibling at school	0.039***	0.020***	0.085***
younger sibling not at school	0.032*	0.266***	0.017
county educational expenditure per capita	-0.986**	0.300	-3.656***
county proportion of pupils aged 15-18 in key schools	-0.018	-0.030	-0.144***
intercept	0.026	0.151***	-0.008
mean value of dependent variable	0.226	0.074	0.352
pseudo R-squared	0.055	0.133	0.095
number of observations	6333	6294	6328

Notes:
1. The dependent variable is 'always performs farmwork after school' = 1; and equivalently for childcare and for housework.
2. The county educational expenditure per capita variable is created by aggregating village educational expenditure per capita to the county level.
3. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.
4. Other variables in the equations but not reported are older sibling working, older sibling not working, older sibling not in school, father's schooling, mother's schooling, county enrolment rate, and county income per capita.

Table 11. The determinants of total household expenditure per enrolled child at primary, middle and high school: OLS estimates

	Primary school	Middle school	High school
ln simulated income per capita	0.011***	0.039***	0.035***
at least one parent a migrant	0.017	0.099**	0.277
older sibling working	-0.011	0.002	-0.525***
older sibling at school	-0.046***	-0.017	-0.396*
younger sibling at school	-0.030***	-0.092***	-0.402***
younger sibling not at school	-0.055***	-0.044	-0.668*
attendance at key school		0.699***	0.622***
minority group	-0.103***	-0.296***	0.010
distance to closest middle school (km)		0.019***	0.047***
county income per capita	0.111***	0.167***	0.668***
county educational expenditure per capita	1.336***	5.197***	15.884**
county proportion of pupils 15-18 in key schools	-0.022	0.223*	1.049
intercept	0.379***	0.483**	1.355***
mean value of dependent variable (000 yuan)	0.426	0.882	2.566
adjusted R-squared	0.252	0.152	0.261
number of observations	3,643	2,555	683

Notes:
1. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.
2. Other variables in the equations but not reported are male, credit constraint applies, and county enrolment rate.

Table 12. The sensitivity of household educational expenditure components to household income and public educational expenditure: OLS estimates

	Primary school	Middle school	High school
ln total expenditure per child			
simulated household income	0.015***	0.018***	0.017**
county income per capita	.0.060***	0.068***	0.125***
educational expenditure by local government	0.178	1.163**	2.665*
ln tuition fee per child			
simulated household income	0.053***	0.019	0.015
county income per capita	0.120***	0.083**	0.168***
educational expenditure by local government	-0.981	5.485*	5.229
ln other expenditure per child			
simulated household income	0.060***	0.031**	0.013
county income per capita	0.130***	0.120***	0.111**
educational expenditure by local government	2.219	2.465	5.664

Notes:
1. Simulated income is income minus income earned by children in the relevant age group (primary 7-12, middle 13-15, high school 16-18 years). County income per capita is the average household income per capita of all sampled households in the county. Educational expenditure by local government is the average village per capita expenditure on education aggregated to the county level.
2. The OLS estimates contain all the explanatory variables of Table 12, but only the coefficients of the three variables of most interest are reported.
3. Since the dependent variables are in logarithmic form, the coefficients represent the percentage increase in expenditure attributable to a unit increase in each explanatory variable.
3. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.

Table 13. Labour allocation of workers among farm, local non-farm, and migration activities: row percentages

	Farming	Non-farming			Ratio of migration to local non-farm
		local	migration	total	
full sample	63.7	22.1	14.2	36.3	64
poorest third of households	75.0	11.7	13.3	25.0	114
richest third of households	51.7	33.8	14.5	48.3	43
poorest third of counties	71.4	12.4	16.2	28.6	131
richest third of counties	54.0	33.8	12.2	46.0	36
poorest half of households in poorest half of counties	73.0	11.9	15.1	27.0	127
workers aged under 30	51.1	18.6	30.4	48.9	163
workers aged under 30 in poorest half of households in poorest half of counties	63.0	9.6	27.5	37.0	287

Note: The criterion for classification between farm and non-farm activity is the predominant number of hours worked, and between local non-farming and migration whether the worker is in or out of the county.

Table 14. Average and marginal returns per hour of labour in farming and non-farming

	Average returns	Marginal returns	
		(1)	(2)
farming	1.44	0.012	0.017
non-farming	2.75	0.143	
of which: local non-farming	3.08		0.167
migration	2.04		0.122
ratio of non-farming to farming	1.91	11.92	
ratio of local non-farming to migration	1.51		1.37

Notes: 1. The marginal returns per hour are based on income functions which contain productive fixed assets and farmland as well as hours worked in the different activities as the explanatory variables
2. All the coefficients in the income functions are significant at the 1 percent level.

Table 15. The determinants of activity choice among farming, local non-farming and migration: multinomial logit analysis

	Local non-farming		Migration	
	coefficient	marginal	coefficient	marginal
male	1.397***	0.183	1.231***	0.067
age	0.084***	0.013	-0.021	-0.003
age squared	-0.001***	-0.0002	-0.001***	-0.0001
college	1.811***	0.355	0.997***	0.025
professional school	1.583***	0.291	1.072***	0.046
high school	0.900***	0.146	0.680***	0.037
middle school	0.657***	0.090	0.528***	0.029
primary school	0.263***	0.035	0.346***	0.023
healthy	0.174***	0.020	0.407***	0.025
farm land (mu)	-0.053***	-0.007	-0.048***	-0.003
county migration density	2.266***	0.198	9.565***	0.672
county non-farm labour density	5.674***	0.796	3.326***	0.162
intercept	-5.263***		-3.044***	
pseudo R-squared	0.284			
mean value of dependent variables	0.221			
number of observations	22,220			

Notes:
1. The activities farming, local non-farming and migration are defined in Table 14. Farming is the reference activity (with coefficients equal to zero).
2. Several explanatory variables were included in the specification but are not reported in the table, including whether the worker has children, has army experience, has suffered a national disaster in the previous year, lives in an officially designated poverty county, etc.
3. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.

Table 16. The determinants of hours worked in non-farm activities by individuals, Tobit estimates

	Coefficients	Marginals	Due to:	
			intensity	participation
male	1099.45***	503.61	171.76	331.85
age	17.30**	7.92	2.70	5.22
age squared	-0.61***	-0.28	-0.10	-0.19
healthy	237.34***	108.71	37.08	71.63
county farm wage	-37.79***	-17.31	-5.90	-11.41
county non-farm wage	74.95***	34.33	11.71	22.62
college	1,412.63***	647.06	220.68	426.38
professional school	1,370.90***	627.95	214.16	413.78
high school	1,038.34***	475.61	162.21	313.40
middle school	724.69***	331.95	113.21	218.74
primary school	284.99***	130.54	44.52	86.02
number of children under 16	-57.30***	-26.25	-8.95	-17.29
number of elderly over 65	-45.87	-21.01	-7.17	-13.84
intercept	-2,264.48***			
mean value of dependent variable	731.74			
pseudo R–squared	0.021			
number of observations	22,172			
number of uncensored observations	10,156			

Notes: 1. The McDonald-Moffitt decomposition is conducted to separate the change in non-farm hours worked into that part due to change in the number of non-farm hours worked (if positive) and that part due to change in the probability of working non-farm hours (if zero).
2. The sample comprises all persons who are currently working and aged under 65 years.

Table 17. The determinants of hours worked in farm activities by individuals, Tobit estimates

	Coefficients	Marginals	Due to:	
			intensity	participation
male	94.12***	80.20	51.58	28.62
age	92.39***	78.72	50.96	28.28
age squared	-0.10***	-0.85	-0.55	-0.31
healthy	35.49**	30.24	19.53	10.84
farm land (mu)	8.56***	7.29	4.71	2.61
productive fixed assets (000 yuan)	10.81***	9.21	5.92	3.29
number of workers	-37.03***	-31.55	-20.30	-11.26
predicted non-farm hours	-0.46***	-0.40	-0.25	-0.14
county non-farm wage	-16.36***	-13.94	-8.98	-4.98
college	-488.48***	-416.22	-267.50	-148.46
professional school	-399.48***	-340.39	-218.74	-121.40
high school	-158.80***	-135.30	-87.05	-48.31
middle school	-133.98***	-114.16	-73.44	-40.76
primary school	-29.27	-24.94	-15.99	-8.87
number of children under 16	26.65***	22.71	14.35	7.96
number of elderly over 65	17.45	14.87	6.31	3.50
intercept	-380.30***			
mean value of dependent variable	955.71			
pseudo R–squared	0.034			
number of observations	22,172			
number of uncensored observations	18,892			

Notes: as for Table 16

Table 18. The determinants of non-farm income of individuals, OLS estimates

	ln non-farm income	ln non-farm income	absolute non farm income (yuan)
male	0.140***	0.086***	305.1***
age	0.053***	0.058***	223.8***
age squared	-0.0006***	-0.0006***	-2.37***
healthy	0.108**	0.111***	213.2*
non-farm hours (instrumented)	0.078***	0.082***	218.0***
college	0.600***	0.734***	4214.0***
professional school	0.425***	0.548***	1986.9***
high school	0.235***	0.319***	1003.0***
middle school	0.173***	0.233***	585.9***
primary school	0.059	0.078	105.3
intercept	4.780***	4.849***	-5,413.4***
county proportion of adults 25-34 who completed high school	0.856		
county mean school fees	0.772***		
county proportion dissatisfied with secondary schools	-0.244		
mean of dependent variable	7.628	7.628	3,780
adjusted R-squared	0.434	0.416	0.274
number of observations	9,513	9,513	9,513

Notes: 1.The standard errors are corrected for clustering at the county level.
2. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.
3. The test for joint significance of the three variables representing quality of education in column 1, F (3, 119) = 11.04, is highly significant.

Table 19. The determinants of farm income of households, OLS estimates

	ln farm income	absolute farm income (yuan)
farm land (mu)	0.042***	245.8***
productive fixed assets	0.041***	257.3***
proportion of workers healthy	-0.024	54.1
(instrumented) hours worked by members with:		
college	0.029*	79.9
professional school	0.047***	272.7***
high school	0.037***	167.8***
middle school	0.031***	141.4***
primary school	0.028***	97.0***
no education	0.017***	19.3*
intercept	7.548***	2,255.0***
mean of dependent variable	8.161	5,375.8
adjusted R-squared	0.183	0.215
number of observations	8,603	8,603

Notes: 1. Farm land is measured in mu, productive fixed assets in 000 yuan, and hours worked in 000 hours.
2. Hours worked are instrumented using the equation reported in Table 17.
3. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.

Table 20. Simulation analysis: the effects on measures of income poverty of improving the education of the labour force

	Actual case	Middle school the minimum	High school the minimum
average farm income	1,275	1,206	1,174
average non-farm income	1,017	1,411	2,069
average other income	432	432	432
average overall income	2,592	3,050	3,675
average farm hours	649	608	564
average non-farm hours	458	555	700
poverty rate (percent):			
$1 a day	10.9	2.5	0.4
$2 a day	43.0	27.0	8.4

Notes: 1. The simulations are based on the equations reported in Tables 16-19.
2. The first simulation assumes that workers without education or only primary school are raised to middle school completion. The second assumes that all workers with less education are raised to high school completion.
3. The $1 and £2 a day criteria are converted into 2002 prices in rural China: 925 yuan and 1,850 yuan respectively.

Table 21. The returns to education in farm, non-farm, and combined activities, by income category

	Farm	Non-farm	Combined
whole sample	0.015***	0.108***	0.064***
poorest third of households	0.006	0.060***	0.020***
richest third of households	0.017**	0.081***	0.045***
poorest third of counties	0.005	0.086***	0.047***
richest third of counties	0.019*	0.068***	0.064***
poorest half of households in poorest half of counties	0.008	0.070***	0.033***
richest half of households in richest half of counties	0.021***	0.079***	0.054***

Notes:
1. The non-farm equations use individual data and the farm and combined equations use household data.
2. In the farm equations the dependent variable is ln household farm income per capita and the explanatory variables are maximum years of education among workers engaged in farming, maximum age among workers engaged in farming and its square, cultivated land (mu) and ln productive fixed assets. The inclusion of the variable input, farm hours, makes very little difference to the estimates of returns.
3. In the non-farm equations the dependent variable is ln individual income and the explanatory variables are years of education, age and age squared. When ln non-farm hours is included the coefficients all fall, reflecting education's influence on the probability and extent of non-farm work.
4. In the combined equations the dependent variable is ln household income per capita and the explanatory variables are maximum years of schooling among workers in the household, maximum age among workers in the household, ln cultivated land (mu), and ln productive fixed assets.
5. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.

Table 22. Returns to education in farming households: OLS estimates

Column:	1	2
average years of schooling	0.042***	0.027***
modern	0.284***	0.140**
modern * average years of schooling		0.021**
land (mu)	0.023***	0.023***
ln productive fixed assets (000 yuan)	0.019**	0.019***
intercept	7.978***	8.084***
mean of dependent variable	8.719	8.719
adjusted R-squared	0.169	0.170
number of observations	4,719	4,719

Notes:
1. 'Farming households' are those that derive a majority of their income from farm activities.
2. 'Modern farming households' are those that derive a majority of their farm income from grain production.
3. ***, **, and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively
4. The dependent variable is ln household farm income.

Table 23. Simulation analysis for the poorest third of households: the effects on measures of income poverty of improving the education of the labour force

	Actual case	Middle school the minimum	High school the minimum
average farm income	769	718	698
average non-farm income	334	443	596
average other income	85	85	85
average overall income	1,072	1,247	1,379
average farm hours	682	635	613
average non-farm hours	262	324	410
poverty rate (percent)			
$1 a day	32.5	21.1	11.7
$2 a day	100.0	96.7	92.2

Notes: As for Table 20, except that the simulations are based on the specifications reported in Table 16-19 but estimated for the sub-sample of households in the lowest third of household income per capita.

Table 24. Simulation analysis for households with heads aged under 30 years: the effects on measures of income poverty of improving the education of the labour force

	Actual case	Middle school the minimum	High school the minimum
average farm income	1,158	1,078	950
average non-farm income	637	1,282	2,387
average other income	440	440	440
average overall income	2,145	2,800	3,777
average farm hours	663	663	513
average non-farm hours	321	397	567
poverty rate (percent)			
$1 a day	16.8	1.8	0.4
$2 a day	52.6	24.7	5.2

Notes: As for Table 20, except that the simulations are based on the specifications reported in Table 16-19 but estimated for the sub-sample of households for which the household head is aged under 30 years.

Table 25. The determinants of health status among adults, binary logit analysis

	Coefficient	Marginal	Mean value	Standard deviation
male	-0.244***	-0.0284	0.5273	0.4993
age in years	0.058***	0.0067	38.21	15.25
age squared	0.00002***	0.0000	1692.61	1296.24
years of schooling	-0.072***	-0.0083	7.0926	2.9733
predicted ln income per capita	-0.192***	-0.0222	7.7706	0.5076
intercept	-1.978***			
pseudo R-squared	0.1518			
proportion unhealthy	17.38			
number of observations	28,729			

Notes:
1. The dependent variable takes a value of one if the respondent reports being so-so, unhealthy or very unhealthy, and of zero if healthy or very healthy.
2. The omitted category in the dummy variable analysis is female.
3. *** denotes significance at the one per cent level.
4. Very similar results were obtained from an ordered probit estimation using all five categories of health status.

Table 26. The determinants of a positive attitude towards new agricultural technology: binary logit analysis

	Coefficient	Marginal
male	0.082	0.017
age	0.074***	0.015
age squared	-0.0008***	-0.0002
years of schooling	0.110***	0.023
ln household income per capita	0.088***	0.018
intercept	-2.167***	
mean of dependent variable	70.00	
pseudo R-squared	0.019	
number of observations	4,725	

Notes:
1. The means and standard deviations of years of schooling are 6.935 and 2.506, and those of ln household income per capita are 8.888 and 0.651 respectively.
2. ***, ** and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.
3. Ln household income per capita is instrumented using productive fixed assets, land, average years of schooling of household workers, number of farm hours worked, number of non-farm hours worked, number of household members, whether the terrain is mountainous, hilly, or plain, and a set of province dummy variables.

Table 27. The determinants of subjective well-being: binary logit and OLS estimates

Happiness measure:	happy or very happy		happiness score
	coefficient	marginal	coefficient
years of schooling	0.020**	0.005	0.006**
ln income per capita (instrumented)	0.703***	0.162	0.278***
poor health	-0.573***	-0.137	-0.237***
current living standard higher	0.742***	0.173	0.307***
current living standard lower	-0.357***	-0.085	-0.256***
age	-0.016	-0.004	-0.012**
age squared	0.0003*	0.0001	0.0002***
male	-0.165***	-0.038	-0.061***
married	0.487***	0.117	0.260***
good mood	1.286***	0.301	0.506***
intercept	-6.464***		-0.023
mean of dependent variable	0.622		2.681
pseudo/adjusted R-squared	0.142		0.212
number of observations	8,861		8,861

Notes:
1. The dependent variable is happy or very happy = 1, so-so, unhappy, or not at all happy = 0 (estimated using binary logit), or a cardinal variable with very happy = 4, happy = 3, so-so = 2, unhappy = 1, not so happy = 0 (estimated using OLS).
2. The independent variables include dummies for current living standard which is higher than five years ago, and current living standard which is lower than five years ago, with the omitted category being current living standard the same; self-reported poor health, with other replies being the omitted category; and self-reported good mood, with other replies being the omitted category.
3. ***, ** and * denote statistical significance at the 1 percent, 5 percent and 10 percent levels respectively.
4. Ln household income per capita is instrumented using productive fixed assets, land, average years of schooling of household workers, number of farm hours worked, number of non-farm hours worked, number of household members, whether the terrain is mountainous, hilly or plain, and a set of province dummy variables