DISCUSSION PAPER SERIES

IZA DP No. 4108

The Gender Education Gap in China: The Power of Water

Yasheng Maimaiti
W. S. Siebert

April 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Gender Education Gap in China: The Power of Water

Yasheng Maimaiti
University of Birmingham

W. S. Siebert
University of Birmingham and IZA

Discussion Paper No. 4108
April 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4108
April 2009

ABSTRACT

The Gender Education Gap in China: The Power of Water

We investigate girls' school dropout rates, bringing forward a novel variable: access to water. We hypothesise that a girl's education suffers when her greater water need for female hygiene purposes after menarche is not met because her household has poor access to water. For testing we use data from rural villages in the China Health and Nutrition Survey. We find that menarche is associated with an increase in the school dropout rate, and indeed the effect is weaker for girls who have good access to water. Water engineering can thus contribute significantly to reducing gender education gaps in rural areas.

JEL Classification: I21, J16, O15, L95, Q25

Keywords: education, gender gaps, menarche, water, China

Corresponding author:

W. S. Siebert
Birmingham Business School
University House
Edgbaston
Birmingham B15 2TT
United Kingdom
E-mail: w.s.siebert@bham.ac.uk

The Gender Education Gap in China - The power of water

> *"If I had to pick the one thing we must do above all else to improve the world, I would say: 'educate girls' "*, (Tharoor 2007: 165).[1]

1. Introduction

In the developing world, the education of girls falls behind that of boys, with widespread evil ramifications, reducing child health and women's earnings (see, e.g., Schultz, 2002), and even women's role within the political system, since there is a strong link between education and political participation (Hannum and Buchmann, 2005). The paper starts from the observation that girls fall behind when the household has poor access to water. Figure 1 demonstrates this fact, using a cross-country dataset. Building on this link, we hypothesise that lack of clean water for girls to wash themselves after menarche is a root cause of the gender education gap.

The argument that lack of water has particular adverse effects on girls' education has been put forward before (below), but data limitations have prevented a test of the magnitude of the effect. We use six waves of the China Health and Nutrition Survey (CHNS), 1989 – 2004, to make the test. The CHNS dataset is well-adapted for this inquiry, because it has information on the household's access to water, as well as detailed information about time spent on household tasks such as water-carrying, and information about the timing of menarche.

We restrict the analysis to children in rural areas, since access to water is primarily a rural problem: according to the CHNS data averaged 1989-2004, over 90% of families in urban China had direct access to tap water compared to only 48% in rural areas. Still, even in 2005 nearly 60% of the Chinese population live in rural areas. Clearly, our investigation will be relevant to women's education in other rural communities in Africa, Asia, and South America to which most of the literature on water and sanitation currently relates.

As we will see, different estimation methods tend to yield similar results which generally support our hypothesis. In most specifications, we find that poor household access to water has a

[1] We are grateful to Charles Diamond, MD of Econostat, for suggesting this quotation. We wish to thank the Carolina Population Center and particularly Shufa Du for advice on the China Health and Nutrition Survey. We are also grateful for valuable suggestions from participants at the University of Amsterdam EALE Conference, the Sheffield Work Pension and Employment Group Conference, the World Bank-IZA Conference in Rabat, Morocco, the Birmingham Business School's Economics workshop, and also Richard Barrett, Marco Ercolani, Bart Golsteyn, John Heywood, Mary O'Mahony, Sylvain Weber, Xiangdong Wei, Yu Zu and Nick Zubanov. Remaining errors are our own.

significant adverse impact on girls' education, and the impact arises after their menarche. As for boys, poor access to water has little impact after controlling for income.

2. The Literature on gender education gaps in China

In China, primary and secondary education takes 12 years to complete, divided into primary, junior secondary and senior secondary stages. In general, primary education lasts 6 years. At junior secondary stage, most have 3 years schooling. The 9-year schooling period in primary and junior secondary schools pertains to 'compulsory' education. General senior secondary education lasts a further 3 years (Yang, 2006). Gender gaps arise particularly in this secondary stage, and the literature can be categorised as follows:

Opportunity cost: According to Li and Tsang (2002), in the past two decades the transition to a market oriented economy has allowed many privately owned enterprises to hire young female workers with limited education in the manufacturing and service sectors, especially in the booming coastal cities. Furthermore, rural villages and towns have developed various small-scale factories and enterprises that hire young women with limited education (see also Connelly and Zheng 2003 and Song et al. 2006). In addition, Knight and Li (1996) argue that girls' education has a higher opportunity cost, since "traditionally" girls are family helpers.

We have detailed information on children's time spent on household work and market work. We will test if working time has different impacts in the education models for girls and boys, and we will also test if the elasticity varies among girls who have good and bad access to water.

Household wealth: this variable is important for children's education, as is parental education and occupation, partly because of their link with wealth and perhaps also because of a link with "values" (e.g., Lauer, 2003) placed on education. Girls' education should arguably be more sensitive to household wealth in the Chinese context of patrilocal marriage traditions (Song et al., 2005). In other words, the woman moves into the husband's family, which therefore benefits from investments in the wife's education, while a son's education directly benefits his own family (see also Li and Tsang, 2002) These arguments imply lower education chances for girls from poor families, which are presumably more sensitive to possible losses due to daughters moving away.

In fact, the literature has not come to a solid conclusion as regards the impact of household wealth on gender gaps in education. Brown and Park (2002) show the importance of household wealth in determining educational outcomes, but there is little indication in their empirical findings that poverty affects girls more severely than boys (see also Connelly and Zheng2003, Song et al 2005, and Yueh 2006). In our analysis we therefore have a twofold interest in the household wealth variable. First, we need to control for household wealth (including parental education and occupation) to ensure that our access to water variable is not simply picking up the poverty of families which have poor access to water. Second, we want to establish whether girls' education is more sensitive to household wealth than boys'.

Siblings: Different numbers of siblings and sibling structure may also have different impacts on parents' education decisions (Conelly and Zheng 2003, Yang 2006, Tsui and Rich 2002). While the findings remain mixed, we agree that detailed controls for sibling structure are necessary.

Field and Arbus's work (2008) on Bangladesh finds that elder sisters experience a greater negative impact of menarche on their schooling than younger. They believe this effect is a result of the Bangladesh custom of early marriage, with elder sisters being required to marry off first after menarche, and their schooling therefore suffering. However, Bangladesh's early marriage rate is much higher than China's. While an estimated 75% of rural girls in Bangladesh (2008: 886) are married before the age of 16, our data give the corresponding figure for China as under 5%. In fact, we derive somewhat similar results, with the menarche being more disruptive for elder sisters' education (given poor access to water – see below). But we believe that this pattern can be more easily explained in terms of elder sisters being the pioneers in combating hygiene problems in difficult environments – younger sisters then taking advantage of this learning.

Geographical location: Children living in remote areas lack nearby schools, adequate transportation and information. All these may have negative impacts on their school enrolment. The negative impact is plausibly larger for girls (Li and Tsang, 2002). However, Connelly and Zheng (2003) find little evidence that living in a hilly county has a significant negative impact on school attendance of girls. We discuss below whether our access to water variable is picking up this location effect – since poor access is likely to overlap with remote geography. Nevertheless, our controls for family income, and also for location will hopefully sweep out this effect.

‘Culture and pro-son bias’: Many researchers address Chinese tradition and cultural factors that may have potential impacts on the school enrolment of boys and girls. Li and Tsang (2002) describe how “families without sons are recorded as having died out”. This rigid lineage system, along with patrilocal marriage patterns might cause a pro-son bias in schooling decisions (though Lee, 2008, finds little evidence of this). Thus, Song et al. (2006) argue that a son’s education is more of an ‘investment’ good in rural China, whereas a daughter’s is often taken as a ‘consumption’ good. Also Knight and Li (1996, 99) argue that “traditionally girls are family helpers” to explain the gender gap.

The difficulty with arguments relying on culture is that we need to know where the “culture” itself comes from (Yueh, 2006). If girls in poor water areas have always had to drop out of school early, a culture will evolve of educating boys, and using the girls as family helpers – or marrying them off (as in Bangladesh). In other words, culture is not exogenous. Clearly, if access to water explains gender gaps, holding family wealth constant, we will not need the concept of pro-son cultural bias.

3. New variables: access to water and menarche

Time costs generated by fetching water are usually put forward as the reason for including access to water as an education determinant, since girls are considered the main water-fetchers in many African and South Asian Countries (Hill and King 1995). At the same time, poor access to water can interact with menarche to reduce girls’ education. Research has found that girls’ health is at risk if proper personal hygiene is not in place after menarche (see Dagwood 1995; Secerino and Moline, 1995), a problem which arises particularly when poor access to water means girls are unable to clean themselves (see Bista, 2004, Nahar, 2006; Kirk and Sommer, 2006; Singh, 1999). Reduced health and cleanliness worries will clearly impact education. However, while economists are beginning to consider the role of the menstrual cycle in economic outcomes (e.g., Ichino and Moretti, 2009, link the cycle to women’s absenteeism), only Field and Arbus (2008)have so far considered the link with education, but without considering the all-important interaction with access to water.

In fact, there is quite a large literature on poor sanitation in rural schools, and the adverse consequences for children’s education (e.g. Behrman et al. 1997), in particular for girls’ schooling (e.g., El-Gilany 2005). Cairncross et al. (1998) find that a school sanitation programme in Bangladesh increased girls’ enrolment by 11 per cent. The problems that can arise for girls are explained in detail by Rose Lidonde (2005) in an African context (seen also Snel and Shordt 2005, and Colclough et al 2000). From this literature we take it that rural schools generally fail to provide good cleaning facilities, and the families of girls have to take this failing into account, and use their own household facilities which are constrained by access to water.

The arguments above boil down to the possibility that that girls' education suffers due to the greater time, health and psychic costs associated with poor access to clean water. Burrows et al (2006, 14) put it well:

> "Lack of adequate water and sanitation both at home and school prevents girls from attending school when menstruating. Girls have a sense of being unclean when there is little clean water to wash themselves, and this can lead them to stay away from school. Also there are rarely private facilities at school where girls can go to the toilet or wash the rags they use during their periods. They can also pick up infections if the water they use to wash rags is dirty, leading to more time off school"

Figure 2 provides a model of our hypothesis. We picture the cost curve as taking a step upwards with menarche in families with poor access to water. For given education demand (D), girls in these families will choose only A years of schooling. Girls in otherwise similar families, but with good access to water – and boys – will choose B years of schooling. Of course, family income will shift both demand and supply curves, and we control for income. Ability is more of a problem, but should not be correlated with menarche or access to water.

Table 1 provides motivation. As can be seen from the first row, girls with good access to water (tap water in the house or courtyard: see Table 2) have a 3 percentage point higher school enrolment rates pre-menarche. However, post-menarche the advantage is larger, 20 points. Using the pre-menarche girls' experience to derive the direct beneficial effects of tap water, we see that tap water raises the enrolment rate of post-menarchal girls by 17 points. The Kaplan-Meier survival curves in Figure 3a also show that girls with poor access to water have a much lower survival curve than the other three groups, with only about a 30% chance of surviving until period 9, which is the end of junior secondary school. As can also be seen, post-menarche girls in poor access to water households have a lower survival probability from the very first school years. They seem to start school late, and progress more slowly. Figure 3b concentrates on the poor access to water households, and shows how boys' survival is similar to the pre-menarche girls. These basic data seem to back up our basic hypothesis, and we will now test it using several multivariate models.

4. The statistical models

To study the interaction of access to water and menarche, we use a difference-in-difference specification:

$y_{it} = X_i\beta_1 + \beta_2 W_i + \beta_3 M_i + \beta_4 W_i \times M_i + \theta_t + \delta_a + \varphi_s + \varepsilon_{ist}$ where i denotes individuals; t denotes time; y_{it} is either years of schooling attained (survival analysis) or an indicator for currently being at school (logit); X_i is a vector of controls including household income, parental occupation/education, the respondent's house and market work, and sibling structure; W_i is an indicator equal to one if the household has no access to tap water (other definitions are also used); M_i is an indicator equal to one if the individual has reached menarche; θ_t is a set of wave dummies; δ_a is a set of age category dummies; and φ_s is a set of county dummies.

β_4 is the coefficient of interest. We expect poor access to water to have a worse impact on girls' schooling after menarche, due to the hygiene related economic and psychological problems they face as described above. These considerations point to a negative interaction between the access to water and menarche variables.

We also control for household income, which is an important variable, as also explained. A further consideration is children's work, both market and household. These variables are endogenous. Fortunately, however, even in rural China few children in practice help much with market work or household work (Table 2).

For initial analyses below we use survival models (Hosmer and Lemeshow, 1999, Cleves et al, 2004), which estimate the probability of dropping out of school conditional on school enrolment until the previous grade. The survival approach has the advantage of using repeated observations on the same individuals where we have them (Table 2 gives details). We use a Weibull model, in which case our estimating equation becomes:

$\ln(y_{it}) = \mathbf{x'\beta} + \ln(\varepsilon_{ist})$, where, *y* is survival time (schooling duration)and $\boldsymbol{\varepsilon}$ is the extreme minimum value distribution with variance or "shape parameter" ρ. Where $\rho > 2$ (see Cleves 2004, 225), as is generally our case, the hazard rate of dropping out of school increases with schooling duration which is reasonable. For any particular variable k, a "time ratio" transformation (Cleves 2004, 209), $\exp(\beta_k)$, is possible, showing the proportionate change in time to failure associated with the kth variable. We will discuss this interpretation.

An alternative is the logit model, based on whether or not the individual is currently at school. We also present results for this specification, using just the first observation for each individual (results are not sensitive to this assumption). Both the logit and the survival models face the difficulty that schooling duration is highly correlated with the pupil's age, and also with menarche, so it is difficult to exert a fine control of age when entering the menarche variable. Hence, the logit model provides a useful alternative view.

A further model we put forward uses as dependent variable the individual's accumulated schooling relative to average schooling for individuals of that age, in that county-year, that is:

$D_{ijta} - \overline{D}_{jta}$ where $\overline{D}_{jta}$ is the average for the a-th age group in county j and year t. We term this dependent variable "age-adjusted schooling". This specification allows good control for age, in that we consider the effects of access to water within exact age groups. The specification also gives a continuous dependent variable which is near normally distributed, allowing simple OLS techniques.

As final model, we use village averages. Here we make up a panel from approximately 100 villages across the six waves. The dependent variable becomes the village average school enrolment rate, and a fixed effect model is possible. Again we use the difference-in-difference specification, testing for a significant negative interaction between the village rate of access to poor water, and the proportion of girls post-menrache in the village schools.

5. Data

The data used in this study come from the Chinese Health and Nutrition Survey (CHNS), jointly conducted by the University of North Carolina and the Chinese Academy of Preventive Medicine, Beijing. The CHNS is designed to examine "how the social and economic transformation of Chinese society and family planning programs implemented by national and local government affect the economic, health and nutritional status of its population" (CHNS, 2007). The survey drew a sample from nine provinces:, Heilongjiang, Liaoning, Henan, Shangdong, Jiangsu, Hubei, Hunan, Guizhou, and Guangxi. As shown in Figure 4, these provinces stretch from the North-East to the South-West, and vary substantially in geography and economic development (Yang, 2006). Households in each province are selected from about 150 rural villages. We group these into 36 counties (normally 4 for each province) in the analysis below.

Our analysis uses information for girls 6-19 from rural areas, where "rural" is defined according to household registration. The restriction to young respondents is necessary because the menarche information is not available for older women (ruling out an analysis of the link between access to water and completed schooling of adults). We do not consider tertiary education, after age 19, since different factors such as marriage enter the schooling decision. In our sample, the mean age of menarche is 12.7 (compare 13.5 in Singh 1999), and 12.9 in El Gilany 2005), and thus, with the 6-19 age span, we have adequate observations both before and after.

The CHNS provides a detailed per capita income estimate for rural households, which is not usually available from other sources. Gross household income in cash or kind is created for different categories and then expenses are deducted to create a net income value, deflated using the appropriate price deflators. To measure income in-kind, the CHNS relies on the respondent's (usually the household head's) estimates of the market value of the goods produced/consumed and received as gifts. For home gardening income, the total value of household food consumed at home or sold is measured. Income from farming, raising livestock/poultry; collective and household fishing; and the value of income from other household business is obtained by same calculations. The CHNS also takes into account welfare subsidies including housing subsidy, child care subsidy and gifts. These data give mean per capita rural household income as 1,225 *yuan* (1988 community CPI) for 1989-2004, which is generally in line with other sources (e.g., 1,067 *yuan* (1990 CPI) for 1987-2001 in Benjamin et al 2005).

In Table 2, the means and standard deviations are shown for variables used in the analysis distinguishing by access to tap water, and also by wave 1989-2004. The top rows give the school enrolment and years of schooling data, and we see how girls only achieve less than boys in households without tap water. In households with tap water there is little gap, with girls in fact doing better. Girls' education has improved faster than boys. The clearest indication of this improvement is shown from the parental education rows lower down the table. Here, we see that the proportion of mothers with high school/college is only 8-16% in 1989, well behind the fathers, but has climbed to 43-44% by 2004, similar to fathers. Thus, China is something of a success story for women's education. Interestingly, girls' education has improved over time, irrespective of access to water: so we must be careful not to claim too much for the "power" of water.

We also see that households generally have more favourable circumstances when they have tap water. Thus, the 1989 measure of household per capita income is 956 yuan in households with tap water, compared to 723 yuan in households without. The difference is even wider in 2004: 2,777 yuan compared to 1,846 yuan. Fathers and mothers are also more educated in households with tap water, though this difference has disappeared by 2004. This improvement in incomes and parental education will explain, at least in part, the earlier onset of menarche over time, since menarche exhibits some sensitivity to nutrition (see Field and Arbus, 2008 for a discussion). As can be seen, the incidence of menarche increases from around half of the girls in the sample in 1989 to nearly two-thirds in 2004.

However, while household income and access to water vary quite closely, there are many villages where rich households have poor access, and poor households have good access. Figure 5 demonstrates this point. We divide households according to whether they are above or below the median for their county, and then show availability of tap water in the household's village. As can be seen, the distributions are bi-modal, with modes at 100% and 0% access to tap water. The distributions differ in the expected direction, with only 23% of poor households living in villages with 100% access to tap water, compared to 45% of rich households. (Another way of making this point is to look at the between village correlation between household income and access to water, which is quite high at –0.37 – Table 9.) Nevertheless, a considerable proportion of rich households live in villages with poor access to water, and vice-versa – and there are many villages where some have good access, and others do not. Therefore, we should have no difficulty in distinguishing the effect of household income on education separately from that of access to water.

Returning to other features of the data, Table 2 gives details of household and market work, and also family structure. We see that girls spend somewhat less time on household and market work in households with good access to water (0.97 to 1.6 hours/day) than in those with poor access (around 1.8 hours/day), presumably because of the convenience that tap water access brings. Of course girls also spend more time on house and market work than boys, which might interfere with girls' schooling. The extra time for girls is about half an hour in households with good access to water, but this time is inconsequential since girls in these households anyway do better than boys as already noted. In households with poor access to water, the difference increases to about an hour, and could have a bearing on the education gap, a fact which we keep in mind in the testing below. Finally, we note a tendency for the single child family structure to become more prevalent over time, whether or not access to water is good.

The last rows of Table 2 show statistics relating to access to water, and the size of the sample. We experiment with different definitions of poor access, sometimes combining water3 and water4, and sometimes taking the worst category, water4, alone to check for sensitivity. As can be seen, there has been an improvement over time in access to water, with the proportion of households with tap water increasing from 33% to 65%. As can be seen, we have 2398 respondents with full data, some of these observed at multiple ages. For the logit analysis we construct a single cross-section, but can use all the data for the survival analysis.

6. Empirical results

Regression results. Table 3 gives results of survival models comparing girls and boys. Here we pool the girls with good and bad access to water, which is not strictly allowable (Table 4), but is useful for description. We find a striking difference in the impact of poor access to water on girls' and boys' schooling. As can be seen, after controlling for all the listed variables, the impact of poor access to water on boys' schooling, while negative (-0.05), is only one-third that for girls (-0.17). These time ratio coefficients mean that poor access to water reduces the duration of girls' schooling by about 16% (=1-0.84), compared to the boys' 5%. As a matter of interest, without controls for household income and parental occupation/education, the effect of the poor water variable for girls approximately doubles in size, as can be seen from the first column. This result illustrates the link between poor household circumstances and access to water which we have already noted from Table 2.

In Table 4, we report results for girls only, testing the effect of the menarche variable. Here we are required to separately distinguish the families with poor access to water from those with better access – a simple difference-in-difference specification is not possible. Still, as can be seen, the effect of menarche is strong for girls in the poor access families, indicating 27% shorter schooling duration. For girls in families with good access the effect is only -15%. These results accord with our hypothesis that menarche has more adverse effects in poor water access households. However, there are other differences between the equations, in particular higher parental occupational status and education seems to have a greater impact on girls in households with good access to water. This finding might imply that such households are better "connected" with the modern economy, and less tradition-bound. Thus, there is still room for an argument that tradition or culture, rather than mere hygiene explains the marked effect of menarche in a poor access to water environment. Hence, we need more evidence.

Table 5 presents some sensitivity tests. Results are shown using different definitions of poor access to water, and also excluding extreme villages - where no households have good access to water (perhaps "tradition-bound" culturally), and where all have access ("modern" villages). Our results might be driven by the cultural contrast between these villages. But, this argument fails, because the pattern of Table 4 generally remains undisturbed, with menarche having two or three times greater impact in households with poor access to water, however defined. The final rows of Table 5 present a different test, based on the suggestion from Field and Arbus (2008), that – if "culture" places the eldest sister in a queue to be married soon after menarche – older sisters schooling should especially suffer (and so should single daughters). Our results show signs that menarche affects younger sisters less, both in the full and poor access to water samples, which backs Field and Arbus up, but the effects are small.

We next take up a logit analysis of school enrollment. Results of two different specifications are given in Table 6, one with all 17 income-related and parental education variables, and the other with clustered income. In both the 'pure' impact of poor access to water disappears after we introduce the menarche variable and the menarche×water cross products. Using clustered income, for example, girls after menarche and who do not have access to tap water are 15% more likely to drop school than those who have access. Interestingly, girls with an older sister do better (9-10% more likely to be at school) which again accords with the Field and Arbus (2008) result. However, in both specifications, the pure effect of menarche remains strong, reducing enrollment by 30%, which is difficult to explain. Probably menarche is picking up the effects of age – though we are controlling for age categories – since older pupils are more likely to drop out.

Next, we specify age-adjusted schooling as the dependent variable. This variable permits better control of age. We now test whether the impact on schooling of poor access to water varies for girls pre- and post-menarche, holding age exactly constant. This control is necessary since age, menarche and school duration are all highly correlated. We need to distinguish our hypothesis from one which asserts that impact of menarche can be interpreted merely as an 'age effect', since it is the older girls who reach menarche first, and are then required to fetch the water.

Table 7 gives the results, and confirms that poor access to water decreases the school attainment of girls after menarche compared to those before, of exactly the same age. We see that having no tap water (water categories 3 and 4) reduces schooling by 0.56 to 1.08 years for girls post-menarche compared to those with tap water, of exactly the same age. However, an interesting result is that menarche now exerts a positive effect, 0.54, on education, presumably because those pupils with earlier menarche are advantaged in unmeasured ways via better family environments. We also introduce some cross products to test, for example, whether the impact of access to water varies by income. Indeed, there are signs that (relative clustered) income helps most (0.34 years) when water access is very poor (water4). We interpret this effect as showing how richer families are better able to buy sanitary pads and cleaning products to protect family members in poor water areas, and so assist schooling.

Summing up over our models, we see that the other variables in the regressions listed in Tables 3– 7 behave much as expected. Table 3 does not suggest much difference between boys and girls, and in particular, lower family income does not disadvantage girls more than boys. In other words, our results do not suggest that lower income households are more "tradition-bound", pushing their daughters out of school and into marriage, while saving their boys. As for the sibling structure variables, we see that having an older brother (Table 4) helps the sister's education more in poor access than in good access to water households, which again does not accord with the view that "pro-son" bias is not important (in line with Lee, 2008). Furthermore, while household work and market work opportunities have significant negative impacts on both girls' and boys' schooling, the impact is not statistically different between boys and girls, and between girls with poor and good access to water.

<u>Magnitudes of access to water effects.</u> Table 8 presents percentiles of survivorship distributions for different scenarios for access to water and family income. Scenarios (1) and (2) show the impact of access to water holding other control variables at their mean, and we see that with poor access, girls at the median lag on average about two years (9.7 – 7.8), whereas the impact of poor access is only 0.7 years for boys (9.5 – 8.7).

In scenarios (3) and (4), we vary household (clustered) income, holding other control variables at their mean. Here we see that being in the poor family category pulls both girls and boys down about the same amount, demonstrating that family income has the similar effects on boys and girls.

In (5) and (7), we assume water access is good and allow the impact of income to vary. We find that girls in fact do better than boys in both the high income (11.2 years at the median vs. 10.7) and low income categories (9.3 years for girls at the median, vs. 9.2). However, assuming water access is bad – scenarios (6) and (8) – girls in both high income and low income categories accumulate about one year's (approximately 15%) less schooling compared to the boys in the same categories. This result again shows how poor access to water poses special problems for girls irrespective of family income.

Policy implications. As a final step, we develop some policy implications using village level water variables. Village level analysis is directly linked to government policy since water engineering works will generally benefit an entire village. We first derive village level variables (about 100 villages per wave) by averaging individual level data. We drop villages for which we have less than 10 observations. Mean values of these variables are given in the first column of Table 9. For policy analysis, we then track changes of the water access variable across the 1989-2004 period, and derive its contribution to the gender gap.

Table 9 presents within- and between-village correlations which give instructive contrasts. The between-village correlations (i.e., simply between village averages) with girls' enrolment rate in the first row are higher, but generally in the same direction, as the within-village correlations in the second column. However, while the rich households variable is well correlated with girls' enrolment between villages (0.47), it presents essentially a zero correlation (-0.02) within villages. The rich households variable is also well correlated negatively between villages with poor access to water (-0.37), but again has zero correlation (0.05) within villages. Thus, within villages, the factors which drive changes in household wealth appear to be independent of the factors which drive changes in access to water and changes in girls' enrolment. A within-village (fixed effect) analysis of girls' enrolment is therefore less likely to confound the benefits of rich households with good access to water.

Table 10 gives the results of the fixed effects (FE) analysis, contrasted with random effects (RE). Results for boys are also presented. As can be seen from the Hausman test in the final row, the FE specification is preferred for girls' education, but makes no difference for boys. Looking down the first column giving the FE results we see that poor access to water by itself has only a small effect on girls' enrolment, 0.07. However, both menarche (-0.17) and particularly the interaction of menarche and poor access (-0.29) reduce enrolment. A one percentage point decline in the village average access to poor water increases girls' school enrolment by about 0.22 (=-0.07+0.29) percentage points holding constant the community rate of girls' menarche. Interestingly, we find that household wealth has no significant effect on girls' enrolment (0.02), other things equal, though it remains significant for boys (0.14)

From these results we can compute the overall contribution of improved access to water on school enrolment. In particular, average girls' school enrolment increased by 14 points (from 62% to 76%) over the 1989-2004 period. Over the same period the proportion of rural households with bad (category 4) access to water fell by 21 points (from 33% to 12%). Therefore, holding other things equal, we can attribute about one-third (=0.22×21/14) of girls' schooling improvement to improved access to water, which is considerable.

Conclusions

Narrowing the gender gap in education is important for economic and social development. Women cannot play their proper role in public affairs unless they are educated as well as men. Democracy itself suffers. In rural China, the gender gap in school enrolment has improved over the past 15 years. However, previous research addressing reasons for the gap and its changes gives mixed results. Our results are reasonably clear: poor access to water appears to have been a definite factor underlying the gender education gap. Moreover, because we have data on menarche, we can point to menarche as the reason why poor access to water has this adverse effect. While our results have problems (chiefly because age, schooling duration and menarche are closely correlated), our different models all point in the same broad direction. Our findings simply underline the fact that periods require water for cleanliness, without which school-going becomes very difficult.

Our research gives estimates of the weight of the disadvantage that girls with poor access to water face. Not having good access to water has little effect on boys, or girls pre-menarche. But, for girls post-menarche, lack of water is a serious disadvantage, causing a 2 years fall in school duration (Table 8), or a 13% fall in the probability of enrolment (Table 6).Our results have a striking policy implication. A major benefit of policies to improve water supplies may not be the obvious household or industrial benefit, but rather an unseen benefit, the improvement in the position of women. While much of these benefits have already been gained in China which has made good progress in raising access to water, our results should be relevant to other areas of the developing world.

Figure 1: Male-female education gap is higher in populations with lower access to water

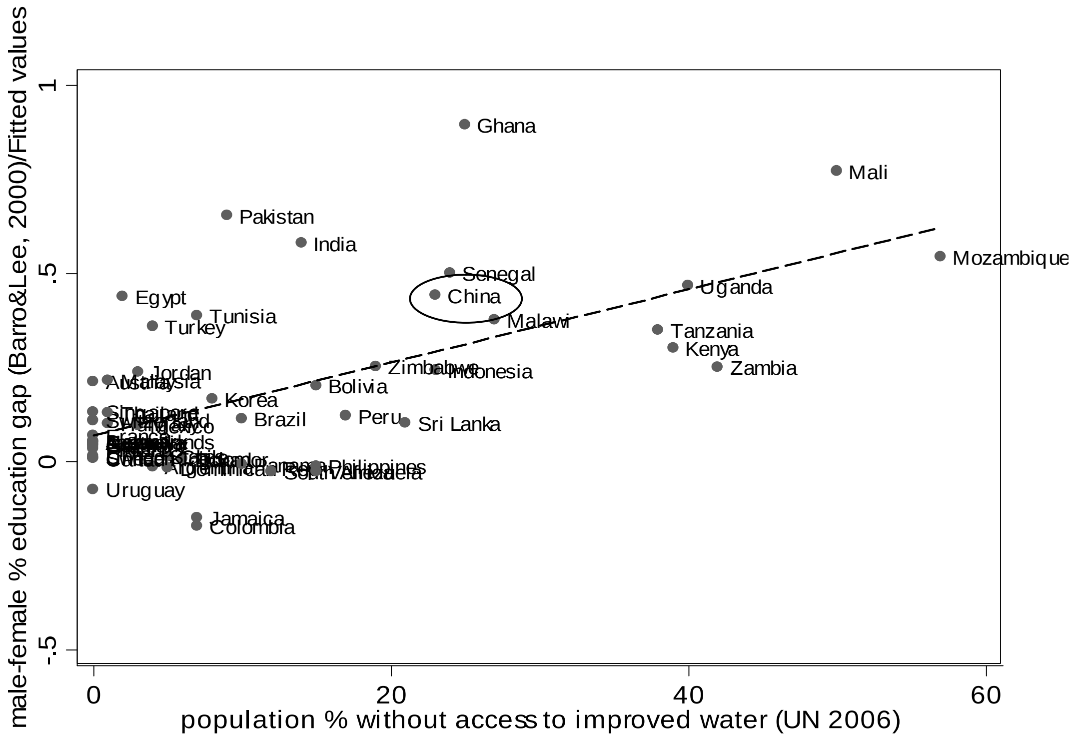


Figure 2: Model of access to water (AW)×menarche interaction and male-female education gap

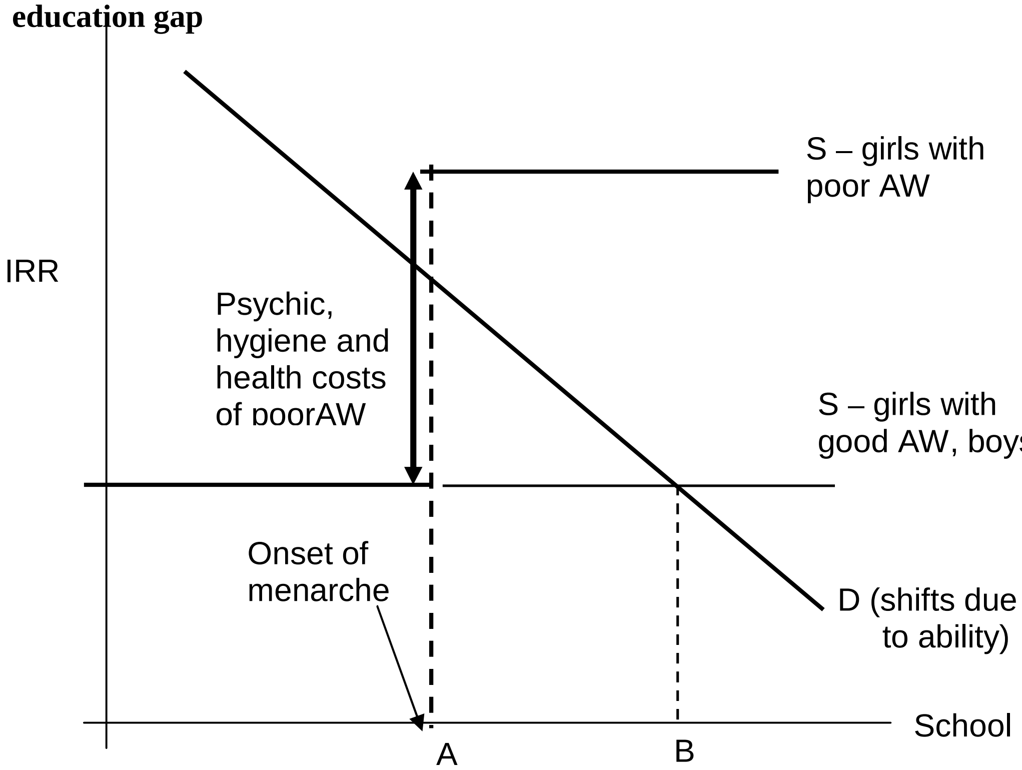


Figure 3a: Kaplan-Meier survival distributions, schooling respondents 6-19, girls only

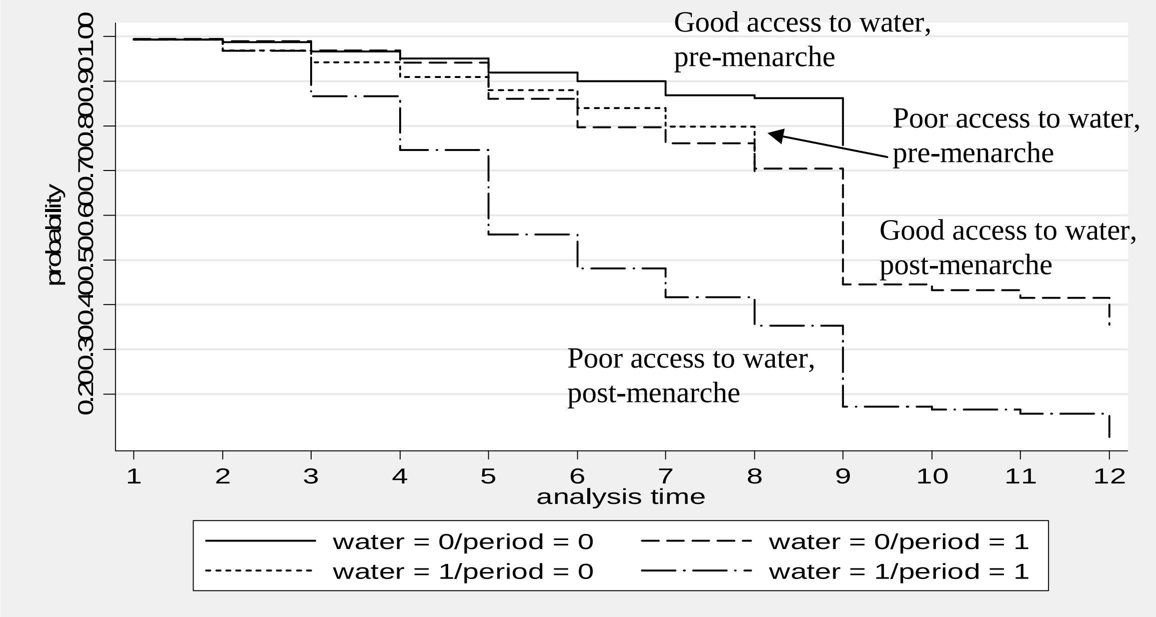
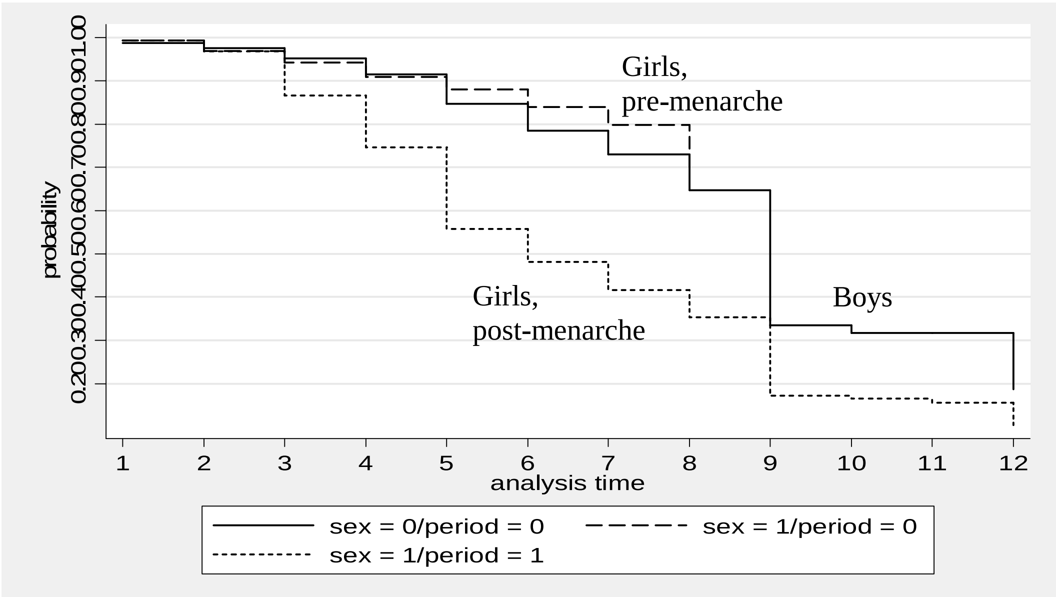


Figure 3b: Kaplan-Meier survival distributions, schooling respondents 6-19 - poor access to water only, boys and girls

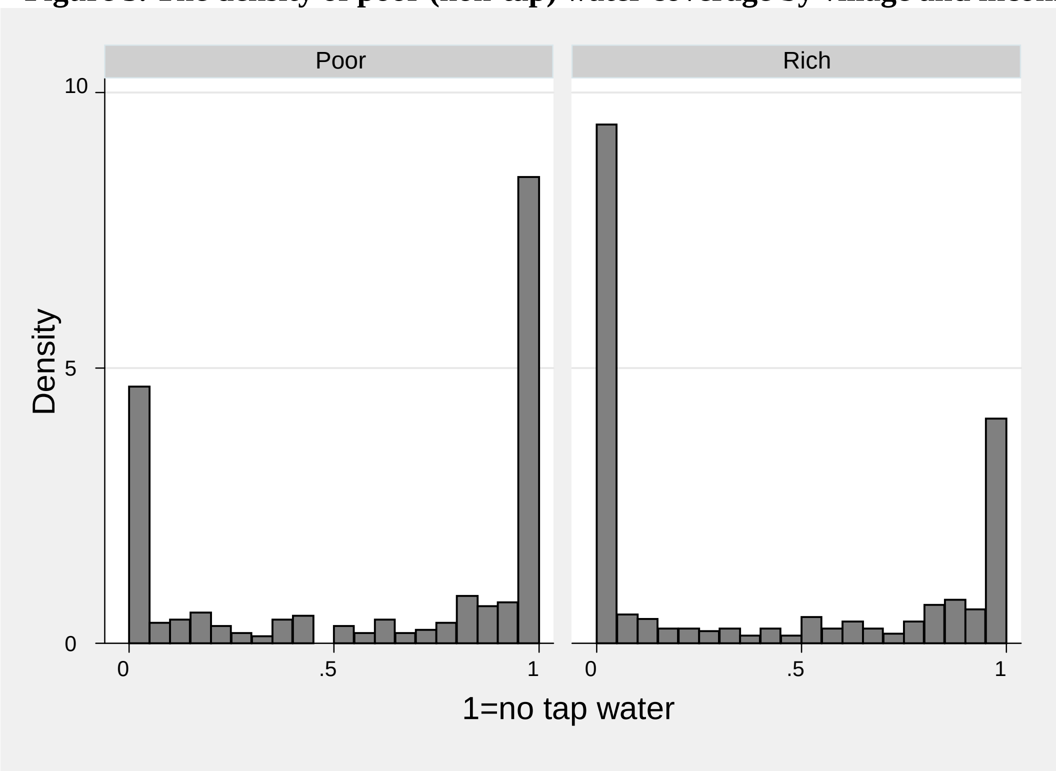


Notes: analysis time corresponds to accumulated years of schooling: 0-6 being primary, 7-9 junior secondary (the end of compulsory education), and 10-12 senior secondary school. Poor access to water is taken as water4 (Table 2).

The reason the survival curve for girls post-menarche extends to low years of education is because some older, post-menarche, girls are in the junior school years.

Figure 4. Map of Provinces Covered in the CHNS Surveys

Figure 5. The density of poor (non-tap) water coverage by village and income category



Note: "poor" and "rich" are identified by the household's per capita income relative to the median for its county. Poor access to water in these figures includes water3 and water4 in Table 2 (non-tap water), and 0 shows 100% of households in a village have access to tap water while 1 shows that no-one has access. The two distributions are tested to be different at the 1% level.

The graphs have 20 bins. As can be seen, about 40% (8*5) of the poor households live in villages where there is no tap water, and about 23% (4.5*5) live in villages where there is tap water for all. The remaining 40% live in mixed access to water villages. As for the rich households, about 45% (9*5) of these live in villages where there is tap water for all, but still about 20% (4*5) live in villages where there is no tap water.

Table 1: Differences-in-Differences Estimate of Poor Access to Water on School Enrolment, Pre- and Post-Menarche

	Girls 6-19		
Menarche	Good access to Water	Poor access to water	Difference
Pre	0.88	0.85	0.03
	(0.014, 614)	(0.013, 788)	(0.018)
Post	0.47	0.27	0.20
	(0.029, 256)	(0.022, 405)	(0.037)
		Difference-in-Difference	-0.17
			(0.036)
Memo: boys 6-19	0.79	0.72	

Notes: Good access to water refers to tap water (categories water1 and water2 in Table 2). Standard error of the estimate and sample sizes are reported in parentheses.

Table 2: Descriptive Statistics, rural girls and boys 6-19 excluding those in tertiary education

Selected Variables	Household has tap water: No 1989	No 2004	Yes 1989	Yes 2004	Description
school enrolment	**0.60**	**0.71**	**0.67**	**0.85**	whether at school (=1) or not (=0)
(girls)	**(0.49)**	**(0.46)**	**(0.47)**	**(0.36)**	
school enrolment	**0.66**	**0.75**	**0.69**	**0.84**	
(boys)	**(0.47)**	**(0.44)**	**(0.46)**	**(0.36)**	
years of schooling	**4.4**	**6.4**	**5.4**	**6.8**	years of schooling
(girls)	**(2.8)**	**(3.0)**	**(3.0)**	**(2.9)**	
years of schooling	**4.8**	**6.8**	**5.3**	**6.6**	
(boys)	**(2.9)**	**(2.9)**	**(3.2)**	**(2.9)**	
menarche	**0.46**	**0.63**	**0.47**	**0.59**	periods begun (yes=1) or no (=0)
	(0.50)	**(0.48)**	**(0.50)**	**(0.49)**	
household per capita income	**723**	**1846**	**956**	**2777**	per capita net real household
(19988 CPI)	**(587)**	**(1491)**	**(554)**	**(2856)**	income (average 1989-2004 is
clustered income*	**0.20**	**0.42**	**0.39**	**0.51**	clustered income (rich=1)
	(0.40)	**(0.49)**	**(0.48)**	**(0.50)**	
father high school/college	**0.21**	**0.45**	**0.24**	**0.42**	
	(0.40)	**(0.49)**	**(0.43)**	**(0.49)**	
mother high school/college	**0.08**	**0.44**	**0.16**	**0.43**	
	(0.28)	**(0.39)**	**(0.36)**	**(0.50)**	
father farmer	**0.71**	**0.47**	**0.46**	**0.32**	
	(0.45)	**(0.50)**	**(0.50)**	**(0.47)**	
household and market work	1.80	1.85	**1.60**	**0.97**	hours spent on household and
hrs/day –girls	(1.90)	(2.16)	**(1.55)**	**(1.36)**	market work per day last week
household and market work	**0.90**	**0.55**	**1.05**	**0.67**	same as above
hrs/day – boys	**(2.02)**	**(1.61)**	**(2.06)**	**(1.89)**	
single child	**0.11**	**0.53**	**0.12**	**0.60**	single child in family
	(0.31)	**(0.50)**	**(0.32)**	**(0.49)**	
Proportion with access to tap water (water1 or water2) (%)	67	37	33	63	4 categories of water are measured: tap water in the home (water1), tap water in the courtyard (water2), well water in the courtyard (water3), or other water outside the courtyard (water4)
Access to tap water or well water in the courtyard (water1, 2, or 3) (%)	33	12	67	88	
Total observations	1,871	302	923	520	
Distribution of the data (1989-2004: 6 waves)	observed in 1 wave only: 40%, 2 waves: 30%, 3 waves: 20%, 4 waves: 8% 5 waves: 2%, all 6 waves: 0 Total girls: 2398 individuals, 6047 observations.				

Notes: Standard Deviations are in parentheses. Bold figures indicate significant differences between 1989 and 2004.
* Clustered income gives two categories, “rich” and “poor”, based on k-mean clustering using whether the household per capita income is above or below median, and mother’s and father’s four job status categories and four educational categories.

Table 3: Schooling Duration (Girls vs. Boys)

Dependent variable: schooling duration (Weibull AFT Models)

	Girls				Boys			
	Time Ratio	z	Time Ratio	z	Time Ratio	z	Time Ratio	z
poor water (water4 - table 2)	0.72***	-7.28	0.84***	-3.65	0.84***	-5.68	0.95	-1.55
log per capita income			1.07***	4.60			1.02	1.48
father - high occ. status			1.22**	2.50			1.09	1.07
father - medium occ. status(ref)								
father - no wage work/farmer			0.92**	-1.97			0.90***	-2.63
father - other/unemployed			0.92	-1.13			0.97	-0.38
mother - high occ. status			1.14	0.83			1.14	1.07
mother - medium occ. status (ref)								
mother - no wage work/farmer			0.81***	-3.55			0.91	-1.54
mother - other/unemployed			0.79***	-3.77			0.89*	-1.80
father - no school qualification (ref)								
father - primary school			1.08**	2.04			1.04	0.97
father - junior middle school			1.12**	2.57			1.14***	3.41
father - high school/college			1.25***	4.00			1.22***	3.69
mother - no school qualification (ref)								
mother - primary school			1.07*	1.63			1.04	1.12
mother - junior middle school			1.06	1.20			1.13***	2.96
mother - high school/college			1.21**	2.40			1.23***	2.91
children's household work			0.96***	-5.43			0.99	-0.22
children's market work			0.91***	-8.86			0.92***	-8.10
single child (ref)								
one older brother			0.92	-1.28			0.88***	-2.82
one older sister			1.02	0.28			1.02	0.41
one younger brother			0.97	-0.49			0.96	-0.63
one younger sister			1.02	0.30			0.97	-0.50
two/more siblings			0.96	-0.74			0.96	-1.03
age group dummies	YES		YES		YES		YES	
county (36) and wave (6) dummies	YES		YES		YES		YES	
				SD				SD
p			3.01	0.15			3.22	0.15
theta			0.26	0.08			0.20	0.09
N				4371				4871
Log likelihood				-1107.3				-1089.2

Note: ρ is the baseline hazard shape parameter (increasing when ρ>1, as here). θ is the frailty or unobserved heterogeneity variance (Hosmer and Lemshow, 1999). **z** scores are calculated using robust standard errors clustered by village. *** denotes significance at 1% level, ** 5% and * 10%.

Table 4: Schooling Duration – Girls only (poor water vs. good water)

Dependent variable: schooling duration (Weibull AFT Models)

	Girls with poor access to water		Girl with good access ♣	
	Time Ratio	z	Time Ratio	z
menarche	0.73**	-2.16	0.85***	-3.18
log per capita income	1.12***	2.78	1.07***	4.38
father - high occ. status	1.31	1.39	1.20**	2.33
father - medium occ. status (ref)				
father - no wage work/farmer	0.90	-0.74	0.93*	-1.80
father - other/unemployed	0.92	-0.41	0.95	-0.71
mother - high occ. status	1.35	0.98	1.14	0.69
mother - medium occ. status (ref)				
mother - no wage work/farmer	1.02	0.08	0.83***	-3.60
mother - other/unemployed	0.87	-0.48	0.81***	-3.59
father - no school qualification (ref)				
father - primary school	1.22***	2.70	1.05	1.23
father - junior middle school	1.30**	2.40	1.05	1.10
father - high school/college	1.10	0.67	1.22***	3.13
mother - no school qualification (ref)				
mother - primary school	1.02	0.18	1.10***	2.61
mother - junior middle school	0.59***	-4.27	1.14***	2.60
mother - high school/college	1.30	0.55	1.24	2.57
children's household work	0.97***	-2.60	0.98**	-2.40
children's market work	0.93***	-3.97	0.93***	-11.00
single child (ref)				
one older brother	1.36	1.54	0.82***	-3.14
one older sister	1.11	0.65	1.01	0.20
one younger brother	1.31	1.50	0.92	-1.48
one younger sister	1.29	1.51	1.01	0.14
two/more siblings	1.15	0.83	0.95	-0.96
age group dummies	YES		YES	
county (36) and wave (6) dummies	YES		YES	
		SD		SD
p	2.43	0.15	3.07	0.13
N	763		3301	
Log likelihood	-221.5		-709.4	

Notes: ♣ poor water is defined as water4 in table 2.

ρ is the baseline hazard shape parameter (increasing when $\rho>1$, as here).

Unobserved heterogeneity variances (θ) are found to be 0 in both specifications (Hosmer and Lemshow, 1999).

z scores are calculated using robust standard errors, and *** denotes significance at the 1% level, ** 5% and * 10%.

Table 5: Coefficient on the Menarche Variable in Different Samples

	Poor water = water3 and water4		
Dependent variable: schooling duration (Weibull AFT Models – Time Ratios)	poor access to water	good access to water	
full sample	0.77***	0.89*	
villages with poor water coverage averaging between 15%-85%	0.60***	0.94	
	Poor water = water4		
full sample	0.73**	0.86***	
villages with poor water coverage averaging between 15%-85%	0.81	0.93	
Dependent variable: at-school=1, 0 otherwise (Marginal Impacts)	The eldest sister	Younger sisters	Single daughter
full sample	-0.49***	-0.40***	-0.44***
water3 and water4 households only	-0.59***	-0.52***	-0.62***

Notes: Controls in all equations are as specified in Table 4. For the logit equations, there are maximum five female siblings in a household and the sample is restricted to households with at least two female siblings in the first two columns.

Table 6: Probability of School Enrolment (Marginal Effects) – Girls only

Logit equations, dependent variable: at-school=1, 0 otherwise

	ME	z	ME	z
poor water (non tap water)	0.01*	0.22	0.02	0.34
menarche	-0.31***	-3.96	-0.30***	-3.74
poor water × menarche	-0.13*	-1.81	-0.15**	-1.98
log income per capita	0.01*	1.85		
father - high occ. status (ref)				
father - medium occ. Status	-0.18***	-2.58		
father - no wage work/farmer	-0.25***	-5.57		
father - other/unemployed	-0.31***	-3.77		
mother - high occ. Status	0.05	0.53		
mother - medium occ. status (ref)				
mother - no wage work/farmer	0.04	0.62		
mother - other/unemployed	-0.02	-0.26		
father - no school qualification				
father - primary school	0.08***	3.40		
father - junior middle school	0.14***	4.09		
father - high school/college (ref)	0.10***	3.92		
mother - no school qualification				
mother - primary school	0.03	1.12		
mother - junior middle school	0.10	1.37		
mother - high school/college (ref)	0.02	0.60		
clustered income			0.08***	3.96
children's household work	-0.04***	-2.90	-0.04***	-3.20
children's market work	-0.07***	-4.39	-0.06***	-4.49
single child (ref)				
one older brother	0.03	0.88	0.01	0.35
one older sister	0.10***	2.83	0.09***	2.78
one younger brother	0.09***	2.78	0.10***	2.98
one younger sister	0.03	0.66	0.03	0.70
two/more siblings	0.07*	1.91	0.04	1.40
agegroup dummies	YES		YES	
county (36)	YES		YES	
Number of observations	2060		2019	
Log likelihood	-772.6		-782.9	

Note: marginal effects are calculated at the means of the independent variables. Robust standard errors account for sample clustering by village. For clustered income, see notes of Table 8. *** denotes significance at the 1% level, ** at 5% and * at 10%.

Table 7: Age-Specific Schooling Analysis, Girls
Dependent variable: age-specific schooling difference

	OLS	
	ME	z
tap water - home (water1) (ref)		
tap water - courtyard (water2)	0.00	-0.02
well water - courtyard (water3)	-0.10	-1.17
other water outside courtyard (water4)	-0.35***	-3.66
menarche	0.54**	2.26
water 2 × menarche	-0.41	-0.93
water 3 × menarche	-0.56*	-1.94
water 4 × menarche	-1.08***	-3.22
clustered income	0.12	1.35
water 2 × clust. income	-0.10	-0.51
water 3 × clust. income	0.18	1.39
water 4 × clust. income	0.34*	1.76
clust. income × menarche	0.16	0.50
water 2 × clust. income × menarche	0.20	0.31
water 3 × clust. income × menarche	0.05	0.11
water 4 × clust. income × menarche	-0.41	-0.75
age_specific household work difference	-0.06	-1.44
age_specific market work difference	0.05	0.69
Number of observations	1744	
Adjusted R^2	0.07	

Note: age-specific schooling difference is $D_{ijta} - \overline{D}_{jta}$ where $\overline{D}_{jta}$ is the average for the a-th age group in county j and wave t – see text. Independent variables are all transformed using the same principle (dichotomous variables will still be dichotomous after the transformation). Marginal effects are calculated at the means of the independent variables. Robust standard errors account for clustering by village. For clustered income, see notes for Table 8. *** denotes significance at the 1% level, ** 5% and * 10%.

Table 8: Percentiles of survival distributions – various scenarios

Survival Time		**0.95**	**0.75**	**0.50**	**0.35**	**0.25**	**0.15**
Good water (1)	Girls	3.9	6.9	9.7	11.7		
	Boys	3.8	6.8	9.5	11.2	12.0	
Poor water (2)	Girls	2.8	5.4	7.8	9.3	10.4	12.0
	Boys	3.7	6.3	8.7	10.3	11.5	
High income (upper cluster) (3)	Girls	4.0	7.5	10.5	12.0		
	Boys	4.2	7.4	10.4	12.0		
Low income (lower cluster) (4)	Girls	3.3	6.5	9.1	10.8	12.0	
	Boys	3.7	6.7	9.3	10.7	11.8	
Good water & high income (5)	Girls	4.4	7.8	11.2			
	Boys	4.3	7.6	10.7			
Poor water & high income (6)	Girls	3.3	6.2	8.7	10.4	11.8	
	Boys	4.0	7.3	9.7	11.4		
Good water & low income (7)	Girls	3.5	6.7	9.3	11.2		
	Boys	3.7	6.5	9.2	10.9	12.0	
Poor water & low income (7)	Girls	2.8	5.2	7.5	8.8	10.0	11.6
	Boys	3.5	6.3	8.6	9.9	11.0	

Notes: Weibull AFT with heterogeneity specification used to create the distributions; other variables in the model are set to their mean values. Poor access to water here includes only category 4 (water outside the courtyard). “Rich” and “poor” categories are based on k-means clustering using per capita household income, parental job status and parental educational qualifications. The cut-off point of per capita household income is set at its median value, while job status and educational qualifications variables are kept in their four original categories. We classify the data into two clusters “poor” (8095 subjects) and “rich” (2303 subjects).

Table 9: Within and between correlations of village level variables

Between Correlations

	means	1)	2)	3)	4)	5)	6)	7)	8)	9)
1) girls' enrolment rate	0.70		**0.49**	**-0.26**	**-0.28**	**0.47**	**-0.70**	**-0.32**	**-0.34**	**-0.31**
2) boys' enrolment rate	0.73	0.27		**-0.16**	**0.01**	**0.26**	**-0.42**	**-0.55**	**-0.36**	**-0.37**
3) proportion with poor access to water	0.18	-0.10	-0.05		**0.05**	**-0.37**	**0.25**	**0.05**	**0.29**	**0.15**
4) proportion of girls post-menarche	0.50	-0.28	-0.02	-0.18		**-0.01**	**0.27**	**0.04**	**0.25**	**0.65**
5) proportion of rich households	0.24	-0.02	0.09	0.05	0.22		**-0.43**	**-0.23**	**-0.32**	**-0.16**
6) work hours for girls/day	0.85	-0.32	-0.05	0.00	0.14	-0.09		**0.50**	**0.35**	**0.42**
7) work hours for boys/day	0.51	-0.08	-0.28	-0.05	0.00	-0.15	0.31		**0.26**	**0.40**
8) number of children per household	2.5	-0.20	-0.14	0.34	-0.12	-0.05	0.07	-0.09		**0.42**
9) average age of children	12.8	-0.21	-0.22	-0.24	0.74	0.17	0.13	0.16	-0.19	

Within Correlations

Notes: within-village correlations are presented in the lower triangle, and between-village correlations are presented in bold in the upper triangle

Table 10: Village school enrolment rates of girls and boys

		Girls		Boys	
		Fixed Effects	Random Effects	Fixed Effects	Random Effects
village rate of poor water (1)		0.07	0.09*	0.01	0.00
		(1.12)	(1.72)	(0.26)	(0.02)
proportion of girls post-menarche (2)		-0.17***	-0.18***		
		(-3.14)	(-3.60)		
(1) × (2)		-0.29**	-0.28***		
		(-2.21)	(-2.72)		
village rate of the rich families		0.02	0.11***	0.14**	0.06**
		(0.34)	(3.85)	(2.37)	(2.30)
village average household work (girls/boys)		0.00	-0.03**	0.00	-0.05
		(0.08)	(-1.98)	(-0.13)	(-1.59)
village average market work (girls/boys)		-0.09***	-0.11***	-0.10***	-0.10***
		(-7.78)	(-10.44)	(-7.50)	(-9.01)
village av. number of children per household		0.01	-0.01	0.02	-0.02
		(0.32)	(-0.70)	(0.48)	(-1.33)
village average age of children		-0.02*	0.001	-0.04***	-0.03***
		(-1.74)	(0.10)	(-6.02)	(-5.47)
wave dummies (6)		YES	YES	YES	YES
R-sq:	within	0.295		0.256	
	between	0.421		0.303	
	overall	0.336	0.371	0.251	0.275
No. of village-years		603	603	603	603
Hausman test – whether FE preferred		Yes, Pr>chi2 =0.087		No, Pr>chi2 =0.139	

Notes: t values are given in parentheses. *** denotes significance at the 1% level, ** at 5% and * at 10%.

References

Behrman, J., R., Khan, S., Ross, D. and Sabot, R., 1997, "School Quality and Cognitive Achievement Production: A Case Study for Rural Pakistan", *Economics of Education Review*, 16 (2): 127-142

Benjamin, D, Brandt, L, Giles, J, Wang, S, 2005, "Income Inequality during China's Economic Transition", working paper, http://www.chass.utoronto.ca/~benjamin/BBGW.pdf

Bista M B, 2004, "A Review of Research Literature on Girls' Education in Nepal", Report prepared for UNESCO Bangkok, http://unesdoc.unesco.org/images/0013/001386/138640e.pdf

Brown, P, and Park, A, 2002, "Education and poverty in rural China", *Economics of Education Review*, 21: 523-541

Burrows, G, J Acton and T Maunder, 2004, Water and Sanitation: the Education Drain, London: Wateraid, http://www.wateraid.org/other/startdownload.asp?DocumentID=113

Cairncross, S., Blumenthal, U., Kelsey, P., Morales, L., and Teeth, A., 1996, "The public and domestic domains in the transmission of disease", *Tropical Medicine and International Health*, 1(1): 27–34.

CHNS, 2007, *China Health and Nutrition Survey*, Chapel Hill, North Carolina: Carolina Population Center, http://www.cpc.unc.edu/projects/china

Cleves M., Gould W. and Gutteriez R., 2004, *An Introduction to Survival Analysis Using Stata*, College Station, Texas: Stata Press.

Colclough, C., Rose, P., Tembon, M., 2000, "Gender inequalities in primary schooling, the role of poverty and adverse cultural practice", *International Journal of Educational Development*, 20: 5-27

Conelly, R, and Zheng, Z, 2003, "Determinants of school enrolment and completion of 10 to 18 year olds in China", *Economics of education review*, 22: 379-388.

Dagwood, M.,Y., 1995, "Dysmenorrhea", *Journal of Reproductive Medicine*, 30(3): 154-167

El-Gilany, A. H., Badawi K., El-Fedawyb, S, 2005, "Menstrual Hygiene among Adolescent Schoolgirls in Mansoura, Egypt", *Reproductive Health Matters*, 13(26): 147-152

Field, E. and Arbus, A., 2008, "Early marriage, age of menarche and female schooling attainment in Bangladesh", *Journal of Political Economy*, 116(5): 881-930.

Hannum E and C Buchmann, 2005, "Global education expansion and socio-economic development: An assessment of findings from the social sciences", *World Development*, 33 (3); 333-354.

Hill, M., A., King, E., M., 1995, "Women's education and economic well-being", *Feminist Economics*, 1(2): 21-46

Hosmer, D, and Lemeshow, S, 1999, *Applied Survival Analysis*, New York: Wiley.

Ichino, A. and Moretti, E., 2009, "Biological Gender Differences, Absenteeism and the Earnings Gap", forthcoming *American Economic Journal: Applied Economics*

Kirk, J. and Sommer, M. 2006, "Menstruation and body awareness: critical issues for girls' education", *EQUALS*, 15: pp 4-5.

Knight, J, and Li, S, 1996, "Educational attainment and the rural-urban divide in China", *Oxford Bulletin of Economics and Statistics*, 58: 83-117

Lauer, C., 2003, "Family Background, cohort and education: a French-German comparison based on a multivariate ordered probit model of education attainment", *Labour Economics*, 10: 231-251.

Lee, Y, 2008, "Do families spend more on boys than girls? Empirical evidence from rural China", *China Economic Review*, 19: 80-100.

Li, D, and Tsang, M.C, 2003, "Household education decisions and implications for gender inequality in education in rural China", *China: An International Journal*, 2003, 1(2): 224-248.

Lidonde, R., 2005, "Why gender is important in SSHE", *WaterLines*, 23(3), 6-9

Nahar, Q, 2006, “Addressing special needs of girls – challenges in school”, Second South Asian Conference on Sanitation, 20-21 September, Islamabad, Pakistan

Schultz, T, 2002, “Why governments should invest more to educate girls”, *World Development*, 30 (2): 207-225.

Secerino S., K., and Moline M.E., 1995, “Premenstrual Syndrome Identification and Management”. *Drugs*. 71-82.

Singh M,.M.,., Devi, R., Gupta S., S., 1999, Awareness and health seeking behaviour of rural adolescent school girls on menstrual and reproductive health problems, *Indian Journal of Medical Science*, 53(10):p 439-443

Snel, M, and Shordt, K, 2005, “The evidence to support hygiene, sanitation and water in schools”, *WaterLines,* 23(3), 2-5

Song, L., and Appleton, S., and Knight, J., 2006, “Why do girls in Rural China have lower school enrolment?” *World Development, 34*(9), 1639_1653

Tharoor, S. 2007, *The Elephant, the Tiger and the Cell Phone: India—The Emerging 21st Century Power*, New York: Arcade Publishing.

Tsui, M, and Rich, L, 2002, “ The only child and educational opportunity for girls in urban China”, *Gender and Society*, 16(1): 74-92

Yang J H, 2006, “Has the One-Child Policy Improved Adolescents’ Educational Wellbeing in China?”, *CHNS Discussion Paper 2006*.

Yueh, L, 2006, “Parental investment in children’s human capital in urban China”, *Applied Economics*, 38: 2089_2111