Reporting Errors, Ability Heterogeneity, and Returns to Schooling in China

Haizheng Li∗
School of Economics
Georgia Institute of Technology
Atlanta, GA 30332-0615
Phone (404)-894-3542, Fax (404) 894-1890
Haizheng.li@econ.gatech.edu

Yi Luo
Kenan-Flagler Business School, CB #3490
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3490
Phone (919)-843-8381
luoy@kenan-flagler.unc.edu

Abstract

A stylized fact in the existing literature is that the economic return to schooling is extremely low in China. Using the newly released household survey data, this study is aimed at providing a more accurate estimate of the returns to schooling in China by controlling for unobserved ability heterogeneity and measurement errors. In addition to employing the commonly used family background variables as instruments, we also identify a unique instrument. Specifically, because boys are favored in the Chinese culture, the presence of any brothers in a family has a negative effect on a girl’s education but presumably has no effect on that girl’s inherent ability. Moreover, efficient GMM estimation is also applied to all models in addition to the 2SLS. We find that the attenuation bias caused by measurement error dominates the omitted ability bias in the OLS estimation. After controlling for both biases, the estimates are considerably higher than the OLS estimates. Based on the GMM estimation, for young workers in China, the return to schooling is 15.0 percent overall and 16.9 percent for women, much higher than the estimates reported from existing studies.

J. E. L Code: I21, J31
Key Words: Returns to schooling, Measurement error, Natural experiment, GMM estimation, China

∗ The corresponding author. An earlier version was presented at the 2002 ASSA meeting in Atlanta. We thank Thomas Dee for discussions and Belton Fleisher for helpful comments.
I. Introduction

A large number of studies have focused on estimating the economic returns to schooling. This literature is generally motivated by a policy debate: should society invest public money in education? An accurate estimate of the effect of education has important policy implications. The traditional approach to estimating the returns to schooling has been to apply the ordinary least squares (OLS) method to estimate the Mincer human capital earnings function (Mincer, 1974; also, see Willis, 1986, for a review). Recent approaches, as reviewed by Card (1999), however, emphasize the so-called “omitted ability bias” problem and apply various instrumental variable (IV) estimations.

In recent years, literature concerning the effect of education on earnings in China has grown. A stylized fact is that the economic returns of education are extremely low in the country. For example, Byron and Manaloto (1990) reported a rate of return of less than 4%; Knight and Song (1991) finds that the effect of education on earnings is remarkably slight; both Johnson and Chow (1997), and Liu (1998) separately estimate the return in the range of 3-4%. A higher rate of return of 5.4% is found by Li (2003), after controlling for heterogeneity in working hours. However, these estimates are still very low, compared to the 10.1% world average and the 9.6% Asian average (Psacharopoulos, 1994).

A common feature of these studies on China’s education is that they followed the traditional OLS approach to estimate the earnings function. Although in modern labor economics there is ample evidence showing a positive correlation between education and earnings, it is still difficult to conclude that the higher earnings observed for better-educated persons are determined solely by their higher education level. The cross-sectional earnings differences among individuals could also reflect inherent ability differences that correlate with education attainment. Thus, the OLS estimation could overstate the true effect of education on earnings. If this is the
case, the estimated returns in China, although very low, may still overestimate the effect of education.

However, in recent literature using IV estimation to correct for omitted ability bias in estimating the effect of education in other countries, published estimates are often substantially higher than OLS estimates (Card 1993, 1995, Butcher and Case 1994, and Ashenfelter and Zimmerman 1997). One possible explanation is that the attenuation bias caused by the measurement error of schooling reduces OLS estimates. Based on Griliches (1977) and Angrist and Krueger (1991), the omitted ability biases in OLS estimates are relatively small, but the downward bias resulting from measurement error can be very large.

In China, the ongoing economic reforms also change the country’s education system. Historically, the government has paid almost all education costs. In recent years, however, the government has started to shift some education costs to individuals. As a result, the tuition costs and fees borne by individuals have increased rapidly, especially for costs covering education beyond the nine-year compulsory schooling level. For example, in 1999 the minimum college tuition was 4,200 yuan per year, which accounted for 72% of urban per capita income of 5,854 yuan and 190% of rural per capita income of 2,210 yuan.1 Facing increasing direct private costs for education, if the earnings premium on education is too low, private demand for education could drop, thus jeopardizing human capital accumulations and economic growth.

Therefore, in order to provide a more accurate estimate of the returns to education, this study attempts to investigate whether the existing estimates based on OLS have biased the true education effect in China; in particular, whether the joint effect of unobserved ability heterogeneity and measurement error underestimates the true returns to schooling, as commonly found in the studies on the casual effect of education in the United States.

---

Since the results of IV estimation depend on the effect of measurement error in schooling, this study will first assess the effect of measurement error; then following the traditional natural experiment approach, use family background variables such as parental education to control for ability bias (Card 1995). Furthermore, we identify a uniquely appropriate instrument from Chinese culture. In particular, because of Chinese families’ cultural preference for boys in a family, the presence of sons in a family results in discrimination against its daughters’ with regard to their education, but presumably has no effect on the daughter’s inherent abilities; thus this presence can be used as an instrumental variable to tackle ability heterogeneity.

In addition, existing studies using IV estimation mostly apply the regular two-stage least squares (2SLS). The 2SLS may be inefficient in many cases, such as when heteroskedasticity is present in the regression errors. In this study, we also apply the Generalized Method of Moments (GMM) procedure in order to get more efficient estimations. Finally, to check the robustness of the results, we conduct formal tests to check the validity of instruments and whether the unobserved ability heterogeneity causes inconsistency in OLS estimation.

The rest of the paper is organized as follows: Section II presents the theory; Section III discuss the data and measurement error; Section IV discuss the results using parental education as control variables vs. as instruments; Section V presents the results uses sibling variables as instruments; Section VI concludes.

**II. Schooling Choice, Earnings, and Ability**

A standard model of schooling choice incorporating ability heterogeneity focuses on the relationship between schooling and average earnings over one’s lifetime (Rosenzweig and Wolpin 2000). Let S denote the level of schooling, \( \alpha \) the ability factor, \( y(S, \alpha) \) the average level of earnings per year an individual will receive if he or she acquires a schooling level S, and \( c(S) \) the direct cost associated with schooling level S when a person completes his or her schooling at \( t=S \). An individual chooses S at \( t=0 \) to maximize the present value of lifetime earnings,
\[ V(S, \alpha) = \int_0^A y(S, \alpha)e^{-rt} dt - c(S)e^{-rS} \]

\[ = y(S, \alpha)e^{-rS}(1 - e^{-r(A-S)})/r - c(S)e^{-rS}, \]

where \( A \) is the maximum working age, and \( r \) is the discount rate.

Suppose that \( y(S, \alpha) \) is an increasing function in \( \alpha \); i.e., given the level of schooling, an individual will receive higher earnings if his or her inherent ability is superior. Then, for a given schooling \( S \), it is likely that there exists a cut-off value of ability \( \alpha^* \), such that individuals with ability at or above that value will acquire schooling level \( S \) and those below it will not; that is, for a given \( S > 0 \), \( \alpha \geq \alpha^* \) is required, where \( y(S, \alpha^*) (1 - e^{-r(A-S)}) = r c(S) \). Therefore, for any specific schooling level \( S \), an individual with greater ability is more likely to complete this schooling level.

For a given ability \( \alpha \geq \alpha^* \), a higher level of schooling is favorable to the individual if

\[ \frac{\partial V(S, \alpha)}{\partial S} > 0. \]

From the first order condition, it is equivalent to

\[ y(S, \alpha) \left( \frac{y'(S, \alpha)}{ry(S, \alpha)} \times \frac{1 - e^{-r(A-S)}}{r} - 1 \right) > c(S) \left( \frac{c'(S)}{c(S)} - r \right). \]

The term \( \frac{y'(S, \alpha)}{y(S, \alpha)} \) measures the marginal returns to schooling, which is \( \beta_1 \) in the conventional Mincer’s human capital earnings function, \( \log y = \beta_0 + \beta_1 S + \beta_2 X + \beta_3 X^2 + \varepsilon \), where \( X \) is the years of work experience. An individual will prefer to attain a higher level of schooling if the marginal return is sufficiently large. Since the marginal returns to schooling is positively correlated with ability, optimizing behavior creates a positive correlation between ability and schooling. Therefore, the differences in earnings among individuals who have attained different levels of schooling will partly reflect these ability differences, which in turn implies that the OLS will overstate the true differences. When using an instrument variable to correct for such bias, the estimated return is expected to be smaller.
However, many studies find that IV estimates are larger than OLS estimates. For example, in Card (1995), and in Ashenfelter and Zimmerman (1997), the use of parental education as an instrument leads to estimates that are at least 15% above the corresponding OLS estimates. In Butcher and Case (1994), IV estimation based on a sibling instrument yields an estimated return of 18 percent, double the OLS result. Card (1993) uses geographic proximity to a four-year college as an instrument for education and also finds that the estimated returns to schooling almost double to 13% from 7%.

Researchers find that such results are caused by measurement error in schooling levels. More specifically, if an individual’s schooling level is measured erroneously and the true value of the returns to schooling is positive, the OLS estimate will be biased toward zero. Thus, the OLS estimate will be too small because of attenuation bias. Based on Card (1999), measurement error bias itself can explain the 10% gap in the estimated returns between OLS and IV estimation.

Therefore, two biases generally exist simultaneously in applying the OLS estimation: the upward bias caused by omitted ability variables and the downward bias caused by measurement error in schooling. If the instruments are not correlated with the measurement error in the schooling level, then IV estimates will be free from both biases. Thus, the result of IV estimation depends on the relative magnitudes of the omitted ability and attenuation biases.

III. Data and Measurement Error in Schooling

The data used in this study are from the second wave of the Chinese Household Income Project (CHIP) conducted in 1996. We use the urban survey, in which 6,928 households and 21,688 individuals in urban areas of eleven provinces were surveyed for 1995 (CHIP-95). The CHIP-95 was funded by the Ford Foundation and a number of other institutes.\(^2\) In the data, annual earnings include regular wages, bonuses, overtime wages, in-kind wages, and other

\(^2\) The CHIP-95 data are available to the public at the Inter-university Consortium for Political and Social Research (ICPSR).
income from the work unit. The hourly wage rate is calculated based on the reported number of working hours. The education measure includes seven degree categories, ranging from below elementary school to college. For more details about the data, see Li (2003).

The sample is restricted to workers aged 30 or below in 1995. This group is selected because the individuals in this group were born in 1965 or after, and in general, attended lower middle school in 1977 or 1978, when the economic reforms started in China. It is possible that educational quality was different prior to economic reform. Before 1977, especially during the Cultural Revolution from 1966 to 1977, the educational system in China deteriorated. Many youths were sent to the countryside for “rectification” (or “re-education”), and colleges and even middle schools were either closed or nonfunctioning. As a result, the quality of education before 1977 may be very different from after then. Thus, we select this group in order to avoid this vintage effect of education that may affect the returns to schooling.

Second, the Chinese educational system was standardized after 1977, and thus it is more accurate to estimate years of schooling based on the degree completed for this age group. In particular, both lower and upper middle school were standardized to three years in length after 1977, whereas before then the length had been two years in some provinces. In addition, because we will use the information on parental education, individuals with such information available are mostly in this age group. The descriptive statistics of the sample are given in Table 1.

As discussed in the previous section, the result of IV estimation depends on the effect of measurement error. In the CHIP-95 urban data, years of schooling are not reported directly, and as in other studies, we estimate years of schooling for each individual based on the degree completed.³ However, this schooling measure could contain errors because the number of years

---

³ In particular, following other studies, the years for a particular schooling level are estimated as follows: for those below elementary school (2 years), finished elementary school (6 years), finished lower middle school (9 years), finished upper middle school (12 years), finished middle level professional or technical school (12 years), finished professional school (i.e., three-year college, 15 years), and finished college degree or above (16 years).
spent obtaining the same degree may vary among individuals. For example, one degree category, “middle level professional or technical school,” can take either two or three years to complete, and can admit graduates from either upper middle level school or lower middle level school. Moreover, another degree category, “college or above,” does not distinguish among bachelors, masters, and doctoral degrees, and thus total years for this category may vary across individuals.

Since it is common to estimate an individual’s years of schooling based on the degree completed in estimating returns to schooling, it is desirable to assess the measurement error. To do so, in general, a second measure of schooling is needed. For example, Ashenfelter and Krueger (1994) obtains the second education measure by asking twins to report on both their own and their respective twin sibling’s schooling level. In the CHIP-95 urban data, since individuals reported years of job experience, the second measure of years of schooling can be obtained using age minus years of working experience minus six, assuming that individuals start school at age seven.\footnote{Generally, in urban China, children start elementary school at age from six to eight. We use starting age of seven here to get the second schooling measure $S_2$. Since $S_2$ is only a rough measure of schooling and key point is that the measurement error in $S_2$ should not be correlated with the measurement error in $S_1$. Thus, the choice of starting age is not essential.} Such an estimate will also contain errors. For example, individuals may have some periods of unemployment/waiting for a job, or may work and attend school at the same time.

To apply the classic theory of errors-in-variables, the measurement errors in two schooling measures should be uncorrelated.\footnote{For a non-classic approach to assessing the effect of measurement error in estimating the effect of education, see a recent study by Kane, Rouse, and Staiger (1999).} Given the sources of errors described, we can assume that the error in estimated years of schooling based on degree categories ($S_1$) is not correlated with the measurement error based on age and job experience ($S_2$). If we write $S_1=S+v_1$ and $S_2=S+v_2$, where $S$ is the true schooling year and $v_i$ ($i=1, 2$) are measurement errors that are uncorrelated with $S$ and with each other, the correlation between the two measures of schooling, $S_1$ and $S_2$, is $\text{Var}(S)/[\text{Var}(S_1)\cdot\text{Var}(S_2)]^{0.5}$. This ratio is sometimes called the “reliability ratio.” In Ashenfelter and Krueger (1994), the correlation of schooling levels reported by the twins is
between 0.88 and 0.92. In our sample, however, the reliability ratio is much lower, at 0.27. It indicates that 73 percent of the measured variance in schooling is error.

In order to assess the attenuation bias in the OLS estimation, we estimate the returns of education using both schooling measures. When S1 is used, the estimated return is 8.9 percent; while when S2 is used, it is 4.4 percent. In general, the schooling measure based on age and experience is less accurate than that based on the degree, and thus has a higher error variance. Therefore, S2 causes a larger attenuation bias and results in a smaller estimated return.\(^6\)

One simple procedure to reduce the effect of measurement error is to use the average of the two schooling measures, because the variance of measurement errors in the average should be smaller. In this case, the estimated return increases to 11 percent, higher than the estimate using either S1 or S2. Such a result indicates that measurement errors exist in S1. More specifically, if there is only measurement error in S2 but no error in S1, then the estimated return based on the average will be attenuated toward zero and should be smaller than that based on S1. Therefore, the commonly used schooling measure based on degree completed contains considerable measurement errors, and will cause attenuation bias in estimating the returns to schooling using the OLS estimation.

IV. Parental Education--Control Variable vs. Instrumental Variable

As shown in Section II, the unobserved ability variables are positively correlated with schooling level, and thus OLS will overestimate the returns to schooling. Generally, there are two approaches to correct for the omitted ability bias: the instrumental variable approach and the control variable approach. The control variable approach is to include proxy variables as additional regressors in the earnings equation to purge or absorb the effect of unobserved ability.

---

\(^6\) It can be shown that for the true return to schooling \(\beta_1\), \(\text{plim}(b_1)=\beta_1 \cdot \sigma_u^2/(\sigma_u^2 + \sigma_v^2)\), where \(b_1\) is the OLS estimate when schooling is measured with error, \(\sigma_v^2\) is the variance of the measurement error, and \(\sigma_u^2\) is the variance of the population error in regressing schooling on other regressors in the earnings equation. Clearly, the higher the variance of the measurement error, the larger the attenuation bias is.
on the relationship between earnings and schooling; while the IV estimation is to instrument the schooling level. For example, some studies use IQ score as a proxy for an individual’s ability (e.g., Altonji and Dunn 1996).

Parental education is a commonly used family background variable. It is still unclear, however, whether parental education should be used as a control variable or as an instrument. Card (1995) and Ashenfelter and Zimmerman (1997) use parental education as a control variable for unobserved ability. The underlying assumption is that parental education may sufficiently influence the extent of their children’s education and would thus be correlated to the capability or productivity of their children at work. Some other studies, however, use parental education as an instrumental variable, assuming parents’ educational levels are not correlated with their children’s inherent abilities but nonetheless are influential on their children’s educational achievements (e.g., Ashenfelter and Zimmerman (1997) use the father’s education as an instrumental as well as a control variable).

These two approaches are based on different assumptions regarding parental education. In particular, the control variable approach assumes that parental education is correlated with an individual’s ability but the IV approach assumes that they are not correlated. When schooling is erroneously measured, an additional difference exists for these two approaches. More specifically, when using control variables, the measurement error in schooling remains intact and the attenuation bias still exists. When using IV estimation, however, both the attenuation bias and the omitted variable bias are corrected, provided that the instruments are not correlated with the measurement error. The net change in the estimated return depends on the relative magnitudes of the two biases. If the attenuation bias is relatively larger, the resulting IV estimate will be higher than the OLS estimate; otherwise, it will be lower.

---

7 It could even exacerbate the attenuation bias if parental education is correlated with the “true” component of schooling level; see Griliches (1977).
We estimate the earnings equation using parental education in both ways. The results are reported in Table 2.\textsuperscript{8} Columns (i) and (ii) illustrate the OLS results. The estimated return based on the OLS without the control variable is 8.9 percent, which is quite high relative to existing studies. The main reason is that our sample focuses on a young group. Based on Li (2003), because of the market oriented economic form, the returns to schooling in urban China are increasing, especially for young people.

When adding parental education as a proxy for omitted ability, the estimated return becomes 7.5 percent, while the estimates for other variables are almost unchanged. This reduction in the estimated return is expected. If parental education is a valid proxy for an individual’s ability, the inclusion of such variables should reduce the upward ability bias and the resulting estimate should be lower. Yet in our result, the difference between the two estimates is small and insignificant. Moreover, it appears that the mother’s education does not have a significant effect on an individual’s wages, while the father’s education is significant; i.e., an additional year of a father’s education increases the individual’s expected wage by 1.9%.

If parental education is a proxy for a child’s inherent ability, then a father’s education and a mother’s education should have similar effects. Therefore, this result calls into question the validity of using parental education as a proxy for a child’s ability. It is possible that the resulting decrease in the estimated return is caused by multicollinearity, because the education of an individual’s parents is positively correlated with such an individual’s schooling level.

When parental education is used as an instrument, the estimated return increases dramatically to 15.6 percent, and is significantly different from the OLS estimate. The magnitude of the increase is similar to the findings in other studies (e.g., Ashenfelter and Zimmerman, 1997). The IV estimation in this case should correct both the attenuation bias and the omitted ability bias because parental education should not be correlated with the measurement error in an

\textsuperscript{8} The sample size becomes smaller due to the availability of parental education information.
individual’s schooling level. It appears that the downward bias caused by measurement error is larger than the upward bias caused by omitted ability variables.

Furthermore, to test whether schooling is correlated with unobserved ability variables in the regression error, a Hausman test is conducted. For simplicity, we apply the regression based Hausman test (Davidson and MacKinnon, 1990). The heteroskedastic robust t statistic is –3.14, and the corresponding P-value is only 0.2%; thus the null hypothesis of exogeneity of schooling is strongly rejected, and OLS estimation is inconsistent.

To examine the validity of using parental education as an instrument, we can check the result from the first stage estimation of the 2SLS. Based on the results, an additional year of paternal education increases the child’s schooling by about 0.18 years, while an additional year of maternal education increases it by 0.13 years; and both are highly significant. Therefore, parental education is indeed correlated with an individual’s schooling level.

Ideally, we want to test whether parental education is correlated with an individual’s ability, i.e., if it is a valid instrument. If we have a subset of valid instruments that identify the model, then we can use the test on over-identifying restrictions to test whether the remaining instruments are valid. With both the father’s and mother’s education available, it is possible to test the over-identifying restrictions. The problem for the test, however, is that we don’t know which instrument is valid because they both come from the same logic.

Nevertheless, such a test can still provide some insight into the validity of parental education as instruments. More specifically, in testing the over-identifying restrictions, we implicitly assume that one instrument is valid (e.g., the father’s education) to test the other (e.g., the mother’s education). If the test rejects the null hypothesis, then the mother’s education is an invalid instrument. If this is the case, then the father’s education should also be invalid because both instruments are chosen in parallel. Our test on over-identifying restrictions follows Basmann (1960). The resulting F-statistic is 0.67 with a P-value of 41%. Thus, the test does not reject the null hypothesis that the additional instrument is valid. Such a result is not against the
use of parental education as instruments, although it does not provide a strong support either. In
the next section, we will return to the over-identification test on parental education using a
different set of instruments.

Furthermore, in the literature of labor economics, an individual’s experience is often
considered to be positively correlated with unobserved motivations and abilities that affect his
wage (Mroz 1987). In studying the returns to education, experience variables represent an
individual’s human capital accumulation through job training (Mincer, 1974). If they are
endogenous, the above IV estimation is still inconsistent. To assess the effect of possible
endogeneity of labor market experience, we use age and age squared to instrument experience
and experience squared. The result is reported in column (iv) of the Table 2. The changes are
relatively small in this case, and the estimated return is 15.3%. Again, the test on over-
identifying restrictions cannot reject the null. Moreover, the Hausman test strongly rejects the
null hypothesis that schooling and experience are exogenous.

Finally, if the regression error is heteroskedastic, as is quite likely in the current cross-
section model, then the 2SLS estimation is inefficient. In order to get a more efficient estimate,
we apply the GMM method (Hansen 1982) to estimate the model. In general, the GMM
estimation is asymptotically more efficient when the regression error is heteroskedastic or serially
correlated (Davidson and MacKinnon 1993, chapter 17). The GMM procedure minimizes a
quadratic form of criterion function based on sample orthogonal conditions between the
regression error and the instruments. The efficient GMM estimation is obtained by using the
optimal weighting matrix, which is the estimated variance-covariance matrix based on the
preliminary GMM estimate. In practice, an iterative procedure is often used in GMM estimation
by replacing the optimal weighting matrix based on the new estimates, in order to improve the
efficiency of GMM estimation in finite sample.

For this model, the convergence is achieved after four iterations with three weighting
matrices. The result is reported in column (v) of Table 2. The GMM estimates are close to the
2SLS results reported in column (iv), and the estimated return is 15.0%. The test on over-identifying restrictions in the GMM framework is different from the Basman procedure. It is based on the minimum value of the GMM criterion function, which is a chi-squared distribution under the null hypothesis. The test cannot reject the null as well.

In the above estimations, we find strong evidence that schooling is correlated with the regression error based on the Hausman test. Therefore, OLS estimation will be inconsistent. Moreover, it appears that the results support the use of parental education as instrumental variables but not as proxies for ability. It is still difficult, however, to verify the validity of parental education as instruments using the over-identification test. In the next section, we will identify new instruments to correct for the omitted ability bias.

V. Sibling Effect and Natural Experiment

Given the difficulty of arguing the validity of parental education as instruments, many studies turn to natural experiments to identify new instruments, such as using data on twins or other sibling variables (e.g., Ashenfelter and Rouse 1998, Butcher and Case 1994). In particular, Butcher and Case (1994) use “the presence of any sisters” within a family as an instrumental variable for the schooling of female workers. They argue that the gender composition of siblings in a family has a significant effect on educational attainment, but no effect on inherent ability, and thus can be used as an instrument.

Following this “natural experiment” approach, we identify a unique instrument based on the Chinese culture. In particular, as a traditional male-dominated society, families place higher values and preferences on having boys in a family. Boys are considered to carry on the family name, while girls are considered to eventually belong to their husbands’ families. Economically, in the absence of a well-established social security system in China, boys will assume the responsibility of taking care of their elderly parents. Thus, having boys serves as a substitute for old age security. Therefore, daughters and sons make different contributions to the life-time
family income (and may have different earnings potential too in China, a country that
discriminations against women may exist). The traditionally strong preference for male children
in China can also be found in family fertility decisions. Based on Zhang (1994), the sexes of a
family’s children have a significant effect on its decision to have more children. For example, the
conditional probability of ceasing to have more children is 24.8% for an average woman with two
daughters. But if either of the daughters were instead a son, the probability would increase to
around 40%.

Just for economic reasons alone, if the parents’ strategy for their children’s educational
investment is to maximize their own expected lifetime income, which is determined by the future
total family income and “security insurance” in their old age, we might expect a systematic
difference in the levels of education obtained by sons and daughters. Therefore, in general, a
family will more strongly emphasize a boy’s education, and a girl’s education may be adversely
affected when she has brothers. For example, when facing financial constraints, a family may
only provide financial support for a boy’s education, while asking a daughter to leave school or
start to work earlier in order to support her brother’s education.

Therefore, it is expected that a girl’s education will be negatively affected by the
presence of any brothers. Naturally, a girl’s inherent ability is not related to whether or not she
has brothers. Thus the presence or the number of brothers can be used as an instrument for
omitted ability in estimating the return to schooling for women. Our sample is restricted to
female workers with 590 observations. A dummy variable indicates the presence of brothers,
and 45% of female workers in the sample have at least one brother. The maximum number of
brothers is 3 in the sample. We first calculate a simple correlation between schooling level and
brother variables. As expected, the correlation coefficient between a woman’s education and the

---

9 The sample is not affected by the "One Child" policy in China. The Chinese government advocated the
“One Couple, One Child” in 1980 for the first time, and officially implemented this policy in 1982.
Individuals born in 1982 or after were thirteen years old or younger in 1995 and were not at working age.
The average number of children raised by one Chinese couple is 4 in the 1970s and 2.4 in the 1980s.
presence of any brothers is –7.6%, and between a woman’s education and the number of brothers in the family is –8.7%; both correlation coefficients are low but highly significant.

In Table 3, for comparison, we first estimate the returns to schooling for women using OLS based on the current sample. The estimated return is 9.8 percent (in column (i)), higher than the overall return given in the previous section. This result is in line with the findings in other studies, i.e., the returns to education is generally higher for women in developing countries as a result of the scarcity of well-educated women (Psacharopoulos 1994).

Since the sibling instrument should not be correlated with the omitted ability variable and the measurement error, we start by using the number of brothers as an instrument because the correlation between women’s schooling and the number of brothers is stronger. The estimated returns to schooling are about 43.8% and are significant around the 10% level. In the first stage of the 2SLS estimation, the number of brothers does have a negative and significant effect on a girl’s education. In particular, the existence of an additional brother will reduce the girls’ education by about 0.3 year.

To check the robustness the result, we also run the same specification for men to see if sisters have a similar effect on boys’ education. The result shows that the existence of sisters does not affect a boy’s education (the coefficient is –0.17 and the corresponding P-value of t-test is 25%). The finding confirms our hypothesis that, in China, the presence of brothers will reduce a girl’s education. But the presence of sisters will not affect a boy’s education.

It is generally desirable to have over-identifying instruments in an IV estimation in order for the IV estimator to have meaningful first and second moments (Kinal 1980). Moreover, additional instruments will increase the asymptotical efficiency of an IV estimator (Davidson and MacKinnon 1993), and the test on over-identifying restrictions can be conducted. Therefore, we use two instruments in order to get over-identified restrictions: the number of brothers and a dummy variable indicating the presence of any brothers. This result is reported in column (ii) of Table 3. The estimated return to schooling becomes 35.6% and is almost significant at the 10%
level. The test on over-identifying restrictions does not reject the null hypothesis. As is known now, if the test rejects the null hypothesis, both instruments will likely be invalid because they are chosen in parallel. Interestingly, the Hausman test on the exogeneity of schooling cannot reject the null hypothesis at the 10% level.

In addition, for women, it is more likely that work experience is correlated with unobservable ability and motivation that affects wages, as commonly found in estimating the labor supply equation for women in the United States (Mroz 1987). To assess the effect of possible endogeneity of experience variables, age and age squared are again used to instrument them. The result is reported in column (iii). The new estimate of the return becomes about 32.6 percent and is significant at the 10% level. The changes in estimates are more sizable than that for the overall sample including both men and women in the last section, indicating that the exogeneity assumption concerning experience has a stronger effect for women than that for men. Again, the test on over-identifying restrictions and the Hausman test do not reject their respective null hypotheses.

One concern for the sibling variables as instruments is that their correlation with schooling is very low. A low correlation can result in a very inefficient IV estimation, especially when the sample size is not very large. Probably, the relatively low efficiency can help explain that the estimated effect of schooling on wages for women is only significant around the 10% level; it may also contribute to the result that the Hausman test cannot reject the exogeneity of the schooling variable. In order to improve efficiency, we also add parental education as an additional instrument. The result is reported in column iv in Table 3. The new estimate of the return to schooling becomes 17.7 percent, and is highly significant. Moreover, the Hausman test rejects the null hypothesis.

In order to further improve the efficiency to account for possible heteroskedasticity, the GMM estimation is also applied to the model. The GMM procedure converges after four iterations with three optimal weighting matrices, and the estimated return is 16.9% (column v).
As discussed in the previous section, an unsolved issue is whether parental education is a valid instrument. The test on over-identifying restrictions there offers an indication but not a conclusion because the instruments (parental education) are chosen based on the same logic. Now, the new instruments, sibling variables, are based on a different logic from that of parental education. Since it is reasonable to assume that the sibling variables are not correlated with a girl’s inherent ability, the subset of sibling instruments can be considered to be valid and they identify the model. Therefore, we can test whether the parental education are valid instruments given the validity of the sibling instruments. The resulting F statistic is 1.03 and is very insignificant (with a P-value of 38%), and thus we cannot reject the null hypothesis that parental education is a valid instrument. The same result is also obtained in the GMM-based over-identification test, with the corresponding Chi-squared statistic of 2.96 and P-value of 40%.

Since parental education appears to be a good instrument, the IV estimation using both sibling instruments and parental education instruments will be more efficient. The resulting IV estimate for women is 17.7 percent from the 2SLS or 16.9 percent from the GMM, almost double the OLS estimate of 9.8 percent. This increase is consistent with the findings in other studies after controlling for both measurement error and omitted ability bias (e.g., Butcher and Case 1994, Ashenfelter and Krueger 1994). The increase also confirms that the effect of measurement error dominates the overall bias in OLS estimation. As a result, after correcting for both attenuation bias and omitted ability bias, the estimated return increases.

VI. Conclusions

This study attempts to provide a more accurate estimate of economic returns to education in China. It is motivated by two stylized findings in the literature. First, existing studies on the effect of education in China find extremely low returns to schooling. Since these studies rely on the OLS procedure, it is natural to ask whether they underestimate the returns. Second, recent studies based on the IV procedure to estimate returns to schooling in the United States generally
find a higher result. It is believed that attenuation bias caused by measurement error dominates the ability bias. The effect caused by possible ability heterogeneity and measurement error has not been studied in the previous research on education in China. This study is aimed at investigating whether these biases have contributed to the very low estimated returns in China.

We use the newly released household survey data and apply various IV estimations to estimate returns to schooling for young workers in urban China. A unique instrument we employed is sibling composition justified by the Chinese cultural preference for boys in a family. Our findings shed new light on the effect of measurement error, omitted ability bias, and the validity of family background variables as control variables or as instrumental variables in estimating the causal effect of education on earnings.

We find that measurement error in schooling causes a considerable downward bias in the OLS estimates. When using IV estimation, the resulting estimates for the returns to schooling are all considerably higher than those from the OLS. Therefore, the attenuation bias caused by measurement error dominates the omitted ability bias. This result is robust using either parental education or sibling variables as instruments.

Based on the GMM estimation, for young workers in China, the estimated returns to schooling are about 15.0 percent overall and 16.9 percent for women, considerably higher than any previous estimates based on the OLS estimation. Such returns are fairly high compared to the Asian average and the world average. The high estimated return obtained in this study can also help to explain why private demand for education (especially at college level and above) is still very strong in China, even though controversies exist on whether the direct private costs of education increase too fast.

Furthermore, we find that in China, the presence of brothers does negatively affect a girl’s education, while the existence of sisters has no effect on a boy’s education. Therefore, sibling variables can be used as instruments for women’s education. In addition, our results do
not support the use of parental education as a control variable, i.e., as a proxy for an individual’s ability. Instead, we cannot reject that parental education can be used as an instrument.

As other studies using the “natural experiment” approach, this study relies on the assumption that a girl’s inherent ability does not depend on her sibling composition. Rosenzweig and Wolpin (2000) discuss the limitation of such an assumption. In addition, due to data constraints, we could only partly assess the effect of measurement error, but cannot separate the attenuation bias.
Reference


Table 1  Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size</th>
<th>Mean Value</th>
<th>Standard Deviation</th>
<th>Minimum Value</th>
<th>Maximum Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage</td>
<td>2511</td>
<td>2.18</td>
<td>1.72</td>
<td>0.0032</td>
<td>29.20</td>
</tr>
<tr>
<td>Schooling 1</td>
<td>2511</td>
<td>12.12</td>
<td>2.26</td>
<td>2.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Schooling 2</td>
<td>2511</td>
<td>12.59</td>
<td>2.57</td>
<td>-3.00</td>
<td>21.00</td>
</tr>
<tr>
<td>Sex</td>
<td>2511</td>
<td>0.48</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Experience</td>
<td>2511</td>
<td>6.66</td>
<td>3.80</td>
<td>1.00</td>
<td>27.00</td>
</tr>
<tr>
<td>Age</td>
<td>2511</td>
<td>25.25</td>
<td>3.37</td>
<td>16.00</td>
<td>30.00</td>
</tr>
<tr>
<td>Ethnic Minority</td>
<td>2511</td>
<td>0.047</td>
<td>0.21</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Father Education</td>
<td>1340</td>
<td>11.029</td>
<td>3.40</td>
<td>2.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Mother Education</td>
<td>1340</td>
<td>9.011</td>
<td>3.45</td>
<td>2.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Presence of Brother</td>
<td>590</td>
<td>0.45</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Number of Male Children</td>
<td>590</td>
<td>0.49</td>
<td>0.57</td>
<td>0.00</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Note: 1. Schooling 1 is estimated by the degree completed; Schooling 2 is estimated by age and years of work experience.
2. “Presence of Brother” is a dummy variable indicating the presence of any brother for a female child in a family.
### Table 2  Parental Education as Control Variables or Instruments

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS (No Control) (i)</th>
<th>OLS (Control) (ii)</th>
<th>IV(1) (iii)</th>
<th>IV(2) (iv)</th>
<th>GMM (v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.091</td>
<td>-1.18</td>
<td>-1.92</td>
<td>-2.039</td>
<td>-1.994</td>
</tr>
<tr>
<td></td>
<td>(-7.70)</td>
<td>(-8.09)</td>
<td>(-6.14)</td>
<td>(-6.91)</td>
<td>(-5.54)</td>
</tr>
<tr>
<td>Schooling</td>
<td>0.089</td>
<td>0.075</td>
<td>0.156</td>
<td>0.153</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>(9.10)</td>
<td>(6.95)</td>
<td>(6.46)</td>
<td>(6.86)</td>
<td>(5.50)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.125</td>
<td>0.127</td>
<td>0.132</td>
<td>0.204</td>
<td>0.203</td>
</tr>
<tr>
<td></td>
<td>(5.09)</td>
<td>(5.18)</td>
<td>(5.22)</td>
<td>(3.50)</td>
<td>(3.41)</td>
</tr>
<tr>
<td>Experience Squared</td>
<td>-0.006</td>
<td>-0.006</td>
<td>-0.007</td>
<td>-0.012</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(-3.69)</td>
<td>(-3.69)</td>
<td>(-3.58)</td>
<td>(-2.84)</td>
<td>(-2.75)</td>
</tr>
<tr>
<td>Sex</td>
<td>0.078</td>
<td>0.078</td>
<td>0.071</td>
<td>0.063</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>(1.78)</td>
<td>(1.78)</td>
<td>(1.59)</td>
<td>(1.37)</td>
<td>(1.32)</td>
</tr>
<tr>
<td>Ethnic Minority</td>
<td>-0.153</td>
<td>-0.156</td>
<td>-0.134</td>
<td>-0.107</td>
<td>-0.112</td>
</tr>
<tr>
<td></td>
<td>(-1.39)</td>
<td>(-1.41)</td>
<td>(-1.18)</td>
<td>(-0.91)</td>
<td>(-0.93)</td>
</tr>
<tr>
<td>Father’s Education</td>
<td>0.019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s Education</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td>1340</td>
<td>1340</td>
<td>1340</td>
<td>1340</td>
<td>1340</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.087</td>
<td>0.094</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Value</td>
<td>26.53</td>
<td>20.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over-identifying</td>
<td>F(1, 1333)=0.67</td>
<td>F(1, 1333)=0.90</td>
<td>χ²₁=0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restriction</td>
<td>P=0.41</td>
<td>P=0.34</td>
<td>P=0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausman Test</td>
<td>t= -3.14</td>
<td>F=5.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.002</td>
<td>P=0.0012</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:**
1. In column (ii), the father’s and mother’s education are used as control variables.
2. In IV(1), the father’s and mother’s education are used as instruments for schooling; in IV(2), father’s education, mother’s education, age and age squared are used as instruments for schooling, experience and experience squared.
3. t-statistics are in parentheses and are calculated based on heteroskedasticity robust standard errors.
4. R² and F statistics are not reported for IV estimation.
5. Because the GMM estimates are close to that from the IV(2), no separate Hausman test is conducted using GMM estimates.
6. The GMM estimation converges after 4 iterations with 3 weighting matrices.
Table 3  Sibling Effect and the Return to Schooling for Women

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS (i)</th>
<th>IV(1) (ii)</th>
<th>IV(2) (iii)</th>
<th>IV(3) (iv)</th>
<th>GMM (v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.260</td>
<td>-4.473</td>
<td>-3.768</td>
<td>-2.213</td>
<td>-2.076</td>
</tr>
<tr>
<td></td>
<td>(-6.03)</td>
<td>(-1.64)</td>
<td>(-1.84)</td>
<td>(-4.97)</td>
<td>(-4.58)</td>
</tr>
<tr>
<td>Schooling</td>
<td>0.098</td>
<td>0.356</td>
<td>0.326</td>
<td>0.177</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>(6.71)</td>
<td>(1.63)</td>
<td>(1.70)</td>
<td>(5.28)</td>
<td>(4.91)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.146</td>
<td>0.173</td>
<td>0.057</td>
<td>0.155</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(3.85)</td>
<td>(3.56)</td>
<td>(0.36)</td>
<td>(1.71)</td>
<td>(1.52)</td>
</tr>
<tr>
<td>Experience Squared</td>
<td>-0.0070</td>
<td>-0.0070</td>
<td>-0.001</td>
<td>-0.008</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(-2.51)</td>
<td>(-2.21)</td>
<td>(-0.11)</td>
<td>(-1.16)</td>
<td>(-0.98)</td>
</tr>
<tr>
<td>Ethnic Minority</td>
<td>-0.494</td>
<td>-0.545</td>
<td>-0.524</td>
<td>-0.515</td>
<td>-0.485</td>
</tr>
<tr>
<td></td>
<td>(-2.73)</td>
<td>(-2.27)</td>
<td>(-2.17)</td>
<td>(-2.65)</td>
<td>(-2.16)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>590</td>
<td>590</td>
<td>590</td>
<td>590</td>
<td>590</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Value</td>
<td>19.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over-identifying Restriction</td>
<td>F(1, 584)=1.63</td>
<td>F(1, 584)=1.67</td>
<td>F(3, 582)=1.03</td>
<td>$\chi^2$=2.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.20</td>
<td>P=0.20</td>
<td>P=0.38</td>
<td>P=0.40</td>
<td></td>
</tr>
<tr>
<td>Hausman Test</td>
<td>t= -1.51</td>
<td>F=0.88</td>
<td>F=2.66</td>
<td>P=0.13</td>
<td>P= 0.45</td>
</tr>
<tr>
<td></td>
<td>P=0.047</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:
1. In IV(1), the presence of any brothers and the number of brothers are used as instruments for schooling; in IV(2), the presence of any brothers, the number of brothers, age and age squared are used as instruments for schooling, experience, experience squared; In IV (3), the presence of any brothers, the number of brothers, age, age squared, the father’s and mother’s education are used as instruments for schooling, experience and experience squared.
2. t-statistics are in parentheses and are calculated based on heteroskedasticity robust standard errors.
3. R² and F statistics are not reported for IV estimation.
4. Because the GMM estimates are close to that from the IV(2), no separate Hausman test is conducted using GMM estimates.
5. The GMM estimation converges after 4 iteration with 3 weighting matrices.