GENDER DIFFERENTIALS IN THE PAYOFF TO SCHOOLING IN RURAL CHINA*

Weiwei Ren
Business School
University of Western Australia

and

Paul W. Miller
School of Economics and Finance
Curtin University

ABSTRACT

This paper examines the gender differential in the payoff to schooling in rural China. The analyses are based on a framework provided by the over education/required education/under education literature, and the decomposition developed by Chiswick and Miller (2008). It shows that the payoff to correctly matched education in rural China is much higher for females than for males. Associated with this, the wage penalty where workers are under qualified in their occupation is greater for females than for males. Over educated females, however, are advantaged compared with their male counterparts. These findings are interpreted using the explanations offered for the gender differential in the payoff to schooling in the growing literature on earnings determination in China.

Keywords: China, Schooling, Earnings, Rates of Return

JEL Codes: J31, J62, J70

* We are grateful to two anonymous referees, and to participants at the 2009 *International Symposium on Contemporary Labour Economics*, at the Wang Yanan Institute for Studies in Economics of Xiamen University, for helpful comments. Ren acknowledges financial support from the University of Western Australia via a University Postgraduate Award and from the Australian Research Council via a Postgraduate Scholarship. Miller acknowledges financial assistance from the Australian Research Council.
Address for correspondence: Paul W. Miller, School of Economics and Finance, Curtin University, GPO Box U1987, Perth, WA 6845, Australia, Email: Paul.Miller@curtin.edu.au, Telephone: +61 8 9266 7758, Fax: +61 8 9266 3026

GENDER DIFFERENTIALS IN THE PAYOFF TO SCHOOLING IN RURAL CHINA

I. INTRODUCTION

An important empirical regularity in studies of the determinants of earnings in China is that the return to education for females exceeds that for males. The studies by Meng (1998) and Li H. (2003) illustrate this clearly. Meng (1998) reported that the payoff to schooling for females in rural China was a statistically significant 2.2 percent, but that for males was a statistically insignificant 1.1 percent. Li H. (2003) found that the payoff to education in urban China was 6.9 percent for females and 4.3 percent for males.

Several explanations have been proposed for this empirical finding. One of these is that it is associated with differences in the demand for, and supply of, education between men and women. Li H. (2003), for example, argues that fewer women achieve high levels of education, which reduces the relative supply of highly skilled women. Li H. (2003) also suggests that the higher return to education for females may be associated with greater positive self-selection of women into the labour force relative to men, whose labour force participation is nearly universal. This more intense positive selection into labour market activity among women would result in the ability of labour market entrants being higher than that of women outside the labour market. The return to education for females is therefore likely to be overestimated in the conventional earnings function, where there is no adjustment for these differences in ability (Zhang *et al.*, 2005). Deolalikar (1993) argues that the gender difference in returns to education is linked to the technology employed in the manufacturing sector, where physical strength is important to productivity. Men have a comparative advantage in physical strength used in unskilled factory positions so that schooling becomes relatively more important and financially rewarding to women who focus on more skill-intensive jobs.

To date, however, there has not been any systematic evaluation of these arguments. In part this is due to the lack of an appropriate framework within which to conduct such an evaluation. In this paper, we use insights from the over education/under education literature (see Hartog, 2000) to investigate the reasons for the differential in the payoff to education between males and females in rural China. This literature proposes that there is a usual or reference level of education for each occupation, and that the earnings of workers will vary according to whether they are correctly matched to that level of education, have more education than the reference level (*i.e.*, are over educated), or have less education than the reference level (*i.e.*, are under educated). Chiswick and Miller (2008) develop a framework within which the gap in the returns to education for two groups can be decomposed into components due to differences in the payoffs to correctly matched education, over education and under education, and to differences in the distribution of workers across these categories. Their decomposition is applied in this paper to account for the gender difference in the payoff to education in China.

The structure of this paper is as follows. Section II provides a brief overview of recent studies of the determinants of earnings in rural China. Section III outlines the methodological framework from Chiswick and Miller (2008). Section IV introduces the data set of the China Health and Nutrition Survey (CHNS), which is used in the empirical section. Sections V and VI present the empirical results. The first of these sections presents results of the regression analyses. Two main models are presented: a conventional schooling and experience earnings function and an Over education/Required education/Under education, or ORU, earnings function. Some sensitivity analyses are also presented. These sensitivity analyses centre on the way the reference level of education is compiled in the ORU model, and on the potential role of sample selection bias. Section VI then undertakes the decomposition of Chiswick and Miller (2008), and presents an assessment of the importance of the path dependence discussed in Chiswick and Miller (2010a) to this decomposition. Section VII concludes.

II. LITERATURE REVIEW

There is now a substantial literature covering the return to education in China. Many of these studies (*e.g.*, Zhao, 1997; Hou, 2004; Yao and Zhang, 2004) have reported that, due to the strict registration of urban and rural workers and the segregation of urban and rural areas with different levels of development, the determinants of earnings differ between workers in rural and urban areas. Accordingly, many studies have undertaken separate analyses for these areas, although most focus on just one area. In particular, most of this research has focused on workers in urban areas, although there are a number of studies of rural workers.[1]

One of the sets of data used for this type of study in rural China is the Chinese Household Income Project (CHIP).[2] The CHIP data for 1988 were, for example, used by Johnson and Chow (1997) to examine the return to education in rural areas as part of a wider study that covered urban workers as well. They reported that in rural areas the returns to education for females and males were 4.82 percent and 2.95 percent, respectively.[3]

Meng's (1998) study is distinguished from the above by its focus only on wage determination in China's rural industrial sector. The data used in her study are from the Township-, Village-, or Privately-Owned Enterprises Sample Survey conducted jointly by the World Bank and the Institute of Economics of the Chinese Academy of Social Sciences in 1986 and 1987.[4] Meng showed that education played slightly different roles in male and female earnings determination. It had a positive (coefficient of 0.022) and statistically significant impact on female wages, but its impact on male wages, while positive (coefficient of 0.011), was statistically insignificant.

A further study of earnings determination in rural areas, by Sun (2002), also reported that the payoff to schooling was higher for females than for males, although the difference was not statistically significant. However, using 2002 data for 15 provinces, which were collected by the research group for the "Study on the relationship between human capital investment and employment of city and countryside in China", Hou (2004) found that the return to education for females was 2.7 percent, but this was less than the 3.9 percent return to education for males in rural areas.

Hence, most studies report that the payoff to schooling for females in rural China typically exceeds that for males. The reasons for these gender differentials in the payoff to schooling are explored in depth below.

III. METHODOLOGY

The approach used in this study is based on Chiswick and Miller's (2008) analysis of the smaller payoff to schooling for immigrants than for the native born in the US. The starting point for their study is the ORU (Over education/Required education/Under education) model of earnings determination, where the natural logarithm of earnings (ln y) is related to the years of required education in the worker's occupation (REQ), any years of over education (OVER) or under education (UNDER) for the worker, labour market experience (EXP) and its square, and other variables, such as location and marital status, that are usually held to affect earnings. That is:

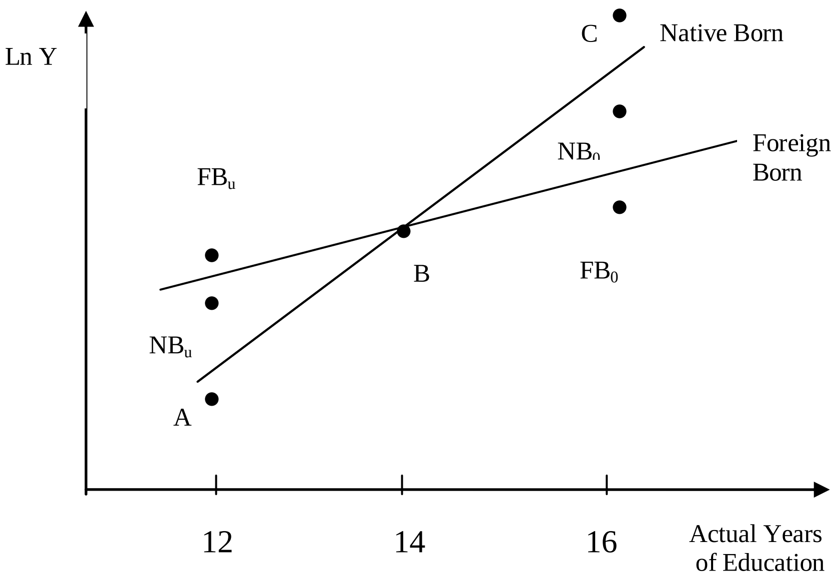
$$\ln y_i = \alpha_0 + \alpha_1 OVER_i + \alpha_2 REQ_i + \alpha_3 UNDER_i + \alpha_4 EXP_i + \alpha_5 EXP_i^2 + ...$$

The way that these differences in the earnings effects of over education and under education impact the return to education in the conventional schooling and experience earnings equation can be illustrated with the diagram from Chiswick and Miller (2008). In this diagram (see Figure 1), the earnings of correctly matched workers with 12, 14 and 16 years of education are depicted by points A, B and C, respectively. As the payoffs to correctly matched schooling are the same for both the foreign born and native born in the Chiswick and Miller (2008) analysis, points A, B and C apply to both birthplace groups.

In the case of over education, which is generally a characteristic of the better educated, the foreign born in the analyses of Chiswick and Miller (2008) had smaller gains associated with 'surplus' education than the native born, and this is represented as FB_0 and NB_0 in the diagram. In contrast, in the case of under education, which will typically be found among the less-well educated, the foreign born at a particular level of schooling had bigger gains associated with 'deficit' education than the native born in the Chiswick and Miller (2008) study, and this is represented by FB_u and NB_u. The conventional or Mincerian returns to education are based around the mean earnings at each level of education. Among the less-well educated, such as those with 12 years of education, these means will comprise the earnings of workers who are correctly matched to the educational requirements of their jobs, where the payoff is the same for the foreign born and native born, and the earnings of under educated workers, among whom the earnings for the foreign born exceed those for the native born. The mean earnings of the less-well educated foreign born will therefore be greater than that for the comparable native born.

Figure 1

Links Between Mincer Payoff to Schooling and Earnings Effects in ORU Model



In contrast, among the better educated, such as those with 16 years of education, while the payoff to correctly matched or required education is the same for the two groups, years of surplus education among the foreign born are rewarded at a lower rate than for the native born. Consequently, the mean earnings of the foreign born at the higher level of education will be below the mean earnings of the native born. These relativities in the mean earnings of the foreign born and native born imply that the earnings-years of the education gradient, or the Mincerian return to education, is less for the foreign born than for the native born. This is illustrated by the two linear lines presented in the Figure.

The discussion with reference to Figure 1 indicates that the reasons for the lower payoff to schooling for the foreign born compared to the native born can be linked back to the earnings effects associated with over education, required education and under education, and the distribution of workers across these categories. Chiswick and Miller (2008) used this information in their decomposition of the difference in the payoff to education for the native born and foreign born as follows.

First, mean predictions of log earnings were obtained at each level of education using the estimated coefficients from the ORU model for the foreign born and the sample values for the ORU variables at each level of education for the same birthplace group. A weighted linear regression that relates these mean earnings to the corresponding years of education was then computed, with the weights being the number of the foreign-born workers at each level of education. This approach mimics the usual calculation of the Mincerian returns to education.

Second, in forming the predictions, the authors again use the foreign-born sample but assume the earnings effects to over education and under education are those estimated for the native born. A weighted linear regression of these mean predicted earnings at each level of education on the years of education was then estimated. Comparison of this implied payoff to schooling with that obtained in the first step above indicates the impact of the differences in the partial effects on earnings associated with over education and under education on the difference in the Mincerian returns to schooling.

Third, when forming the predicted earnings at each level of schooling, it is assumed that there is the same extent of over education and under education within each schooling category for the foreign born as for the native born. Again, a weighted linear regression of these predictions against the levels of schooling was estimated to provide an indicator of the payoff to schooling under the equal returns, equal distributions across over education, required education and under education categories assumptions. Comparison of this payoff with that derived in the previous step shows the role of the different distributions of workers across the over education/under education categories to the difference in the Mincerian payoff to schooling.

Fourth, the linear regression considered in the previous step was re-estimated using the distribution of workers across schooling categories for the native born as the weights. Comparison of the payoffs obtained here with those in the previous step shows the role of the disproportionate representation of the foreign born among the lower education categories where under education is prevalent.

Chiswick and Miller (2008) linked the over education and under education in the ORU model to aspects of the migration process and immigrant adjustment. In particular, under education was linked to positive selection in immigration, particularly among immigrants with low levels of schooling. Over education was linked to the less-than-perfect international transferability of immigrants' country-of-origin human capital. Similarly, in this study the under education and over education phenomena will be linked to the explanations for the higher payoff to schooling for females in China that have been proposed in the literature.

IV. DATA

The data used in this paper are drawn from the China Health and Nutrition Survey (CHNS). The CHNS is a collaborative project of the National Institute of Nutrition and Food Safety, the Chinese Centre for Disease Control and Prevention, and the University of North Carolina at Chapel Hill. The survey was conducted in 1989, 1991, 1993, 1997, 2000, 2004, and 2006 and covered 9 provinces, namely Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning and Shandong. These provinces vary substantially in geography, economic development, public resources, and health indicators. Four counties were selected in each province.

In addition, the provincial capital and a lower income city were selected when feasible. The survey took place over a three-day period using a multistage, random cluster design to select a sample of about 4400 households with a total of 19000 individuals.

Because of the focus on the returns to education in rural China in this paper, only individuals in rural areas with positive wage and salary earnings are included in the analysis; those who are retired, in school, or working part-time are excluded. Owners of private or individual enterprises have also been excluded, because it is difficult to separate their wages from profit income. The potential for the restriction of the sample to workers in paid employment in rural areas to result in sample selection bias is considered as part of a general approach to this potential problem in the sensitivity analyses of Section V. Observations with missing values on education, experience, etc. have been dropped.

The CHNS has detailed information on years of schooling for most individuals. The only possible shortcoming of the data on educational attainment arises in assigning a particular number of years of schooling to the category "6 years college/university or more", which includes both workers with master's degrees and doctoral degrees. This phenomenon will cause errors in estimating the years of schooling, as well as in estimating years of potential labour market experience, although the small number of workers involved means this source of error should not be overly important.

In this study, the earnings measure is just monthly regular wages, excluding the earnings of secondary jobs, all kinds of subsidies and annual bonuses, as the latter have the potential to depend on group productivity to a greater extent than they depend on individual productivity.[5] While the data set contains information on secondary jobs, only around 15 percent of the sample reported such positions, and there is earnings data for only a small subset of these. Owing to this severe missing data problem, it is necessary to focus only on primary jobs.

Finally, as the review of the literature reveals that the gender differential in the payoff to schooling has been observed for all time periods, only data from a recent (2006) wave of the survey are used. After deleting observations with missing data from this wave, 979 observations remain, of which 614 are males and 365 are females.

The CHNS is also used to compile the usual or reference level of education in the ORU model. There are three methods through which this has been established in the literature (see Hartog, 2000), namely: Job Analysis, which is a systematic evaluation by professional job analysts who specify the required education for the occupation titles in an occupational classification; Worker Self-assessment, where workers specify either the education required for their jobs or whether they effectively utilise their levels of education in their work; and Realised Matches, which is based on either the mean or modal level of education in the workers' occupations. Only the last measure is feasible with the CHNS. Thus, the main set of analyses below is conducted using the modal value of education as the required level, and the robustness of the results is examined using the mean level of education of the workers' occupation.[6]

V. REGRESSION RESULTS

The first sub-section below presents estimates of the conventional schooling and experience earnings function. The second sub-section then covers both the incidence of over education and under education in China's rural areas and the empirical results based on the ORU earnings function. In each sub-section, particular emphasis is placed on the differences between males and females.

Table 1 provides the definitions of the variables included in the estimating equations. Means of these variables for the male and female samples are presented in Appendix A.

(i) Mincerian equation results

Table 2 presents the estimated coefficients obtained from applying a standard schooling and experience model to the CHNS for 2006. Estimates are displayed for all workers (Columns i and ii) and separately for males (Column iii) and females (Column iv).

Table 1

Definition of Variables

Variable	Definition
Dependent Variable	
Log Earnings	Natural logarithm of monthly income.
Independent Variables	
Years of education	Actual years of schooling, which is entered into the estimating equation as a continuous variable.
Required education	Years of required schooling based on the modal value of education for every occupation. This is a continuous variable.
Over education	Years of over education, which is the number of years of schooling in excess of the required level of education.
Under education	Years of under education, which is the number of years of schooling less than the required level education.
Log hours	Natural logarithm of working hours in each week.
Experience	Years of potential labour market experience, measured as: age-schooling-6.
Female	Dichotomous variable: Female=1 if female, and Female=0 if male.
Married	Dichotomous variable: Married=1 if married, and Married=0 otherwise.
Ruralreg	Dichotomous variable: Ruralreg=1 if rural registration, and Ruralreg=0 if urban registration.
Central [(1)]	Dichotomous variable: Central=1 if located in the central areas, and Central=0 otherwise.
East	Dichotomous variable: East=1 if located in the Eastern areas and East=0 otherwise.
Stat[(2)]	Dichotomous variable for ownership: Stat=1 if the individual works in a government department, state service/institute or state-owned enterprise, and Stat=0 if otherwise.
Prov	Dichotomous variable for ownership: Prov=1 if the individual works in a private enterprise such as family contract farming, private, individual enterprise or three-capital enterprise (owned by foreigners, overseas Chinese and joint ventures), and Prov=0 if otherwise.
Clerk[(3)]	Dichotomous variable for occupation: Clerk=1 if the individual works as an office staff (secretary, office helper) or an ordinary soldier, policeman, and Clerk=0 if otherwise.
Junior	Dichotomous variable for occupation: Junior=1 if the individual works as a junior professional/technical worker, skilled worker (foreman, group leader, craftsman) or driver, and Junior=0 if otherwise.
Senior	Dichotomous variable for occupation: Senior=1 if the individual works as a senior professional/technical worker (doctor, professor, lawyer, architect, engineer), and Senior=0 if otherwise.
Leader	Dichotomous variable for occupation: Leader=1 if the individual works as an administrator/executive/manager (working proprietor, government official, section chief, department director, administrative cadre, village leader) or an army officer, police officer, and Leader=0 if otherwise.
Other	Dichotomous variable for occupation: Other=1 if the worker's occupation is unknown, and Other=0 if otherwise.

Notes: [(1)] West is the benchmark region; [(2)] Collective enterprise is the benchmark ownership; [(3)] The benchmark occupation is unskilled workers (ordinary labourers, service workers, farm workers).

Table 2

Estimates of Mincerian Model of Earnings Determination in Rural Areas

Variables	All Workers		Males	Females
	(i)	(ii)	(iii)	(iv)
Constant	4.7287*** (16.19)	4.7796*** (16.76)	4.7711*** (12.33)	4.3079*** (9.86)
Years of education	0.0524*** (6.90)	0.0413*** (4.96)	0.0381*** (3.68)	0.0718*** (6.39)
Log hours	0.2958*** (4.33)	0.2927*** (4.36)	0.3059*** (3.60)	0.3210*** (2.86)
Experience (Exp)	0.0268*** (4.07)	0.0230*** (3.48)	0.0240*** (2.81)	0.0253*** (2.70)
$Exp^2/100$	-0.0446*** (3.52)	-0.0383*** (3.02)	-0.0466*** (2.76)	-0.0277 (1.52)
Female	-0.2374*** (6.33)	-0.2223*** (5.82)		
Married	-0.0458 (0.72)	-0.0259 (0.42)	0.0599 (0.81)	-0.1239 (1.23)
Ruralreg	-0.1474*** (3.17)	-0.1478*** (3.20)	-0.1784*** (3.05)	-0.1374* (1.86)
Central	0.0556 (1.17)	0.0284 (0.60)	0.1514*** (2.59)	-0.1433* (1.88)
East	0.1868*** (4.54)	0.1713*** (4.11)	0.2381*** (4.74)	0.1261* (1.89)
Stat	0.1334** (2.03)	0.0952 (1.45)	0.0987 (1.23)	0.2110* (1.87)
Prov	0.1272** (2.07)	0.1544** (2.35)	0.1909** (2.52)	0.0312 (0.30)
Clerk		0.1560** (2.28)		
Senior		0.2849*** (3.68)		
Junior		0.2023*** (4.21)		
Leader		0.1321 (1.19)		
Other		0.1876 (1.63)		
Adj. R^2	0.1795	0.1947	0.1307	0.2085
Sample size	979	979	614	365

Notes: Numbers in parentheses are heteroscedasticity-consistent '*t*' statistics; * denotes that the variables are significant at the 10% level; ** denotes that the variables are significant at the 5% level; *** denotes that the variables are significant at the 1% level. See the Appendix for definitions of variables.
Source: China Health and Nutrition Survey (CHNS), 2006.

There are four features of Column (i). First, the coefficient on the years of education variable shows that the return to an additional year of education is 5.24 percent. Second, the estimated earnings-experience profile displays all the usual features. The payoff to a year of experience is quite high at low levels of experience (*e.g.*, it is 2.23 percent at 5 years) and lower at high levels of experience (*e.g.*, 0.90 percent at 20 years). The earnings-experience profile peaks at 30 years of experience, which is consistent with previous studies. Third, according to the coefficient on the gender dummy variable, females earn 21.1 percent less than males.[7] Fourth, the coefficients of other variables are similar to findings reported elsewhere, with the possible exception of the marriage variable, which is statistically insignificant. In part this appears to be due to pooling the data across two samples (males and females) where the marriage variable has opposite effects. However, even in the analyses conducted on the separate samples of males and females, the estimated effects for the marriage variable are imprecisely determined.

Column (ii) of Table 1 augments the Mincerian model with five occupation dummy variables. The inclusion of these controls for the occupation of employment is associated with an increase in the adjusted R^2, from 0.1795 to 0.1947. The estimates for the occupation variables show that earnings differ by about 30 percentage points across the five occupations distinguished in the analysis, being lowest in the benchmark group of unskilled workers and highest in the senior group which comprises professional/technical workers (doctors, professors, lawyers, architects and engineers). It is found that after controlling for the occupational structure, the return to education falls, from 5.24 percent to 4.13 percent, a 21 percent decline. That is, 21 percent of the increase in earnings related to additional education occurs through entrance into higher-paying occupation. The remaining portion of the return to education is associated with higher earnings within the major group occupations.

The equations estimated on the separate samples of males and females reveal that the return to education for males is 3.81 percent (Column iii). In contrast, the partial effect of years of schooling on earnings for females is much higher, 7.18 percent (Column iv). The gender difference in the payoff to schooling is thus about 3.37 percentage points, which is a similar finding to that reported in the other studies on China covered in Section II. The gender difference in returns to education in China established in Columns (iii) and (iv) above is the focus of the remainder of this section.

(ii) Over education and under education in rural China

Utilising the Realised Matches method, Table 3 presents the distributions of individuals across the correctly matched, over educated and under educated categories in rural areas in China in 2006.

Table 2

The Distribution (%) of Workers across the Correctly Matched, Over Educated and Under Educated Categories in Rural Areas

Variables	All workers (i)	Males (ii)	Females (iii)
Correctly matched	45.7	34.9	49.1
Over educated	27.3	37.5	24.9
Under educated	27.0	27.7	26.0
Sample size	*979*	*614*	*365*

Source: China Health and Nutrition Survey (CHNS), 2006.

There are two main features of Table 3. First, the percentage of correctly matched workers is 45.7 percent. The percentages of over educated and under educated workers are, respectively, 27.3 percent and 27.0 percent. Second, the distribution across the correctly matched, over educated and under educated categories for males is very different from that for females. Hence, the difference in the mismatch for males and females in these data indicates that analysis using the ORU model may be important. The effects in this regard will depend on the coefficients of the variables in the ORU model of earnings determination. Table 4 lists estimates of the ORU model that explore this issue.

Columns (i) and (ii) of Table 4 present the results from the ORU for the total sample. The first specification is for the simple model that does not control for occupation of employment. The second specification is distinguished by the inclusion of five dummy variables for occupation of employment. Columns (iii) and (iv) present the results from the ORU model (without variables for occupation) for the separate samples of males and females, respectively. All of the ORU variables that are the distinguishing feature of this set of results are statistically significant in the column (i) specification. Moreover, the adjusted R^2 in the ORU is higher than in the conventional earnings equation. Comparing Column (i) in Table 1 and Column (i) in Table 4, we see that the coefficients of other variables, such as experience, marital status and regions, are similar between the Mincerian model and the ORU specification. Accordingly, the discussion of Table 4 can focus on the estimated effects associated with the three ORU variables.

Table 4

Estimates of ORU Model of Earnings Determination in Rural Areas

Variables	All workers (i)	All workers (ii)	Males (iii)	Females (iv)
Constant	4.4080*** (12.76)	4.7341*** (12.53)	4.4010*** (9.99)	4.1136*** (7.20)
Reference education	0.0771*** (4.23)	0.0409* (1.74)	0.0635*** (3.18)	0.0935** (2.44)
Over education	0.0637*** (4.91)	0.0559*** (4.22)	0.0578*** (3.55)	0.0583*** (2.88)
Under education	-0.0345*** (2.58)	-0.0296** (2.25)	-0.0075 (0.39)	-0.0763*** (4.16)
Log hours	0.3052*** (4.47)	0.2949*** (4.38)	0.3169*** (3.71)	0.3232*** (2.88)
Experience (Exp)	0.0302*** (4.34)	0.0258*** (3.64)	0.0296*** (3.23)	0.0242** (2.42)
$Exp^2/100$	-0.0528*** (3.80)	-0.0446*** (3.16)	-0.0593*** (3.23)	-0.0241 (1.14)
Female	-0.2431*** (6.44)	-0.2235*** (5.64)		
Married	-0.0568 (0.89)	-0.0350 (0.56)	0.0387 (0.51)	-0.1280 (1.29)
Ruralreg	-0.1474*** (3.17)	-0.1492*** (3.23)	-0.1808*** (3.10)	-0.1301* (1.77)
Central	0.0502 (1.05)	0.0331 (0.70)	0.1406** (2.42)	-0.1599** (2.06)
East	0.1880*** (4.57)	0.1750*** (4.21)	0.2422*** (4.84)	0.1193* (1.76)
Stat	0.1034 (1.52)	0.0965 (1.43)	0.0762 (0.94)	0.1796 (1.43)
Prov	0.1445** (2.26)	0.1560** (2.36)	0.2159*** (2.76)	0.0412 (0.38)
Clerk		0.1587* (1.86)		
Senior		0.2861*** (3.14)		
Junior		0.2000*** (3.93)		
Leader		0.1278 (1.09)		
Other		0.1860 (1.61)		
Adj. R^2	0.1819	0.1945	0.1364	0.2060
Sample size	979	979	614	365

Notes: Numbers in parentheses are heteroscedasticity-consistent '*t*' statistics; * denotes that the variables are significant at the 10% level; ** denotes that the variables are significant at the 5% level; *** denotes that the variables are significant at the 1% level. See the Appendix for definitions of variables.

Source: China Health and Nutrition Survey (CHNS), 2006.

From Column (i) in Table 4, we can readily see the three typical features of the ORU specification of the earnings function. First, the return to the required level of education is 7.71 percent, and this is higher than the 5.24 percent return to actual education (see Table 2, Column i). Second, the return to over education is 6.37 percent, which is almost 1.3 percentage points lower than the return to required education. Third, the coefficient on the under education variable is -3.45 percent. This shows that under educated workers earn 3.45 percent less than adequately educated workers per year of under education.

Column (ii) of Table 4 augments the ORU specification (Column i) with five dummy variables to control for occupation of employment. As a result of this change the return to required education falls, from 7.71 percent to 4.09 percent, a 45 percent decline. Moreover, the required education variable is no longer statistically significant in the aggregate-level analyses. The explanation for this change is similar to that advanced when discussing the Mincerian model. Without occupation variables, the return to the reference years of education includes the effect of moving to an occupation where the schooling can be most effectively used as well as the effect of schooling on earnings within the existing occupation. Once the occupation variables are controlled for, the worker mobility is constrained to be within the broad occupational groups distinguished in this analysis. It is this constraint on worker mobility that is the reason why the return to reference education in the ORU specification falls when the occupation variables are included in the model.

The results of the ORU model for males and females in Columns (iii) and (iv) reveal that there are several gender differences in the estimated impacts of the ORU variables. For males, the return on the required level of education is 6.35 percent, 2.5 percentage points more than that obtained when the actual years of education variable is used in the specification. For females, the return to the required level of education is much higher, 9.35 percent, which is 2.2 percentage points more than the return to actual years of education, of 7.18 percent. This suggests that a female worker with the same actual years of schooling who is correctly matched to the educational requirements of their job can earn three percentage points more per year of schooling than a correctly matched male worker when other things are equal.

The return to over education for males is 5.78 percent, and this is similar to the 5.83 percent return for females. However, because the return to the required education for females is 9.35 percent, around three percentage points greater than that for males, the loss associated with over education for females is still greater than that for males. For instance, compared to a correctly matched worker, a worker who is over educated by one year would be worse off by 3.5 percentage points if female and by 0.6 percentage point if male.

Years of under education are associated with a small, and statistically insignificant, earnings penalty of 0.75 percent for males, and a large, and statistically significant, earnings penalty of 7.63 percent for females. That is, the loss associated with under education for males is very minor while that for females is quite pronounced.

In order to test the robustness of the results, the mean level of education in each occupation was also used as the required level of education in the ORU model (results not reported here but available from the authors upon request). When the required level of education is changed from the mode to the mean, the regression results change only a little. The coefficient on required education based on the mean is 8.90 percent for females, which is a little less than one-half a percentage point below the 9.35 percent based on the mode. The coefficient on over education for females is 5.21 percent based on the mean, a little less than the 5.83 percent based on the mode. The coefficient on under education for females changes from -7.86 percent based on the mode to -7.63 percent based on the mean. Similar small changes to the coefficients of the ORU variables are observed for males as the mode is replaced by the mean value. Thus the conclusion can be drawn that the regression results are not sensitive to the measuring base, which is consistent with other studies (*e.g.*, Hartog, 2000, Chiswick and Miller, 2010b).

A potential shortcoming of the estimates of the earnings equation presented above is sample selection bias. There are two potential sources of selection bias, deriving, respectively, from the labour force participation decision and the choice of area of employment. Several selection correction frameworks can be considered in this situation (such as the use of two independent probit equations: see Choudhury (1994); and the multinomial logit selection model developed by Lee (1983), with non-participation being one alternative, along with area of employment—see Miller (2009) and the references therein). A multinomial logit model is used in the current sensitivity analysis.

An important consideration with this application is the identification of the selection effect. Two approaches have been taken for identification, namely the use of variables that affect participation in the paid labour force or the choice of area but not wages, and relying upon functional form considerations. Included in the latter are the non-linearity of the lambda terms, and the use of different representations of key variables, such as educational attainment and age/experience, in the selection equation and the wage equation. For example, education might be entered in the wage equation as "years of education" and it might be entered in the multinomial logit selection equation as dummies for the various levels of achievement (see, for example, Hartog and Oosterbeek, 1993; Gyourko and Tracy, 1988). Relying on the non-linearity of the lambda selection term is generally viewed as offering a weak means of identification, and using different functional forms for variables in the selection and wage equations is generally argued to involve arbitrary choices. In many empirical applications of the sample selection correction, the results appear to be very sensitive to the specific approaches taken (see Puhani (2000) for discussion), and it appears that the analyses can be more sensitive to the distributional assumptions inherent in the conventional sample selection corrections than they are to the omitted variables bias associated with failure to deal with the sample selection issue (Miller, 1987).

Functional form (the non-linearity of the lambda term) was used for identification in the first instance to obtain a full set of regressions with a correction for sample selection (different functional forms for the education and labour market experience variables were also considered, and the findings are noted below).[8] However, the sample selection term was significant in only one equation. Estimates of the lambda term and it's 't' statistic are presented in Table 5. The full set of results is available from the authors upon request.

Table 5

Estimates of Coefficient on Sample Selection Term in Rural Areas, Mincer and ORU Models

Sample and Model	All Workers		Males	Females
	(i)	(ii)	(iii)	(iv)
Mincer	0.2039 (0.80)	0.2493 (0.97)	0.8026** (2.07)	0.0656 (0.18)
ORU	0.1741 (1.25)	0.2206 (1.59)	0.2380 (1.26)	0.1981 (0.92)

Notes: 't' statistics are presented in parentheses. ** = significant at the 5 percent level.
Source: Statistical Appendix, available upon request.

It is apparent from Table 5 that there is limited evidence in favour of sample selection bias being an important issue. Re-estimation of the selection equation using higher order terms (cubics and quartics) for the education and experience variables as an additional form of identification did not alter this conclusion. Hence the remainder of this study is based on the Ordinary Least Squares (OLS) findings.

Thus, according to the ORU model based on both the mode as the reference level of education and on the mean as the reference level, estimated using OLS, there are significant differences between the earnings of workers in the under educated, correctly matched and over educated categories. These differences can be illustrated by considering the five types of workers described in Table 6. For this explanation, the monthly earnings of Type B workers have been set to 892 yuan, which is the mean of earnings of the total sample in rural areas pooled across males and females. Types A, B and C workers are all matched to the level of education required in their occupation, but they are different in terms of actual years of education.

Table 6

Earnings of Hypothetical Workers in Rural Areas

Type	Actual years of education	Reference years	ORU classification	Hypothetical earnings	
				Males	Females
A	9	9	Correct match	737	674
B	12	12	Correct match	892	892
C	16	16	Correct match	1150	1296
U	9	12	Under educated	872	709
O	16	12	Over educated	1124	1126

Compared to Type B workers, Type A workers and Type C workers have three fewer and four more years of required and actual education, respectively. Then based on the return to reference years of education, the monthly earnings of Type A workers and Type C workers can be predicted as 737 yuan and 1150 yuan, respectively, for males.[9] The monthly earnings of Type A female workers and Type C female workers are predicted as 674 yuan and 1296 yuan, respectively.[10]

Type U workers have three years of actual education less than the level of education required in their occupation. That is, they are under educated. Based on the return to under education, the monthly earnings of Type U male and female workers can be predicted as 872 yuan and 709 yuan, respectively.[11] Similarly, we can predict the monthly earnings of over educated, Type O, workers as 1124 yuan for males and 1126 yuan for females.[12]

Figure 2 displays the monthly earnings of these five types of workers in rural areas. It also includes hypothetical earnings-years of education relationships derived from these monthly earnings, which illustrate the higher payoff to education for females than for males. The decomposition analyses presented in the next section quantify the roles of over education, correctly matched education and under education in generating these different earnings-years of schooling profiles for males and females.

VI. DECOMPOSITION ANALYSIS

(i) Basic Results

Results from application of the Chiswick and Miller (2008) decomposition to the results from the estimation of the ORU earnings model for rural China are given in Table 7.

Figure 2

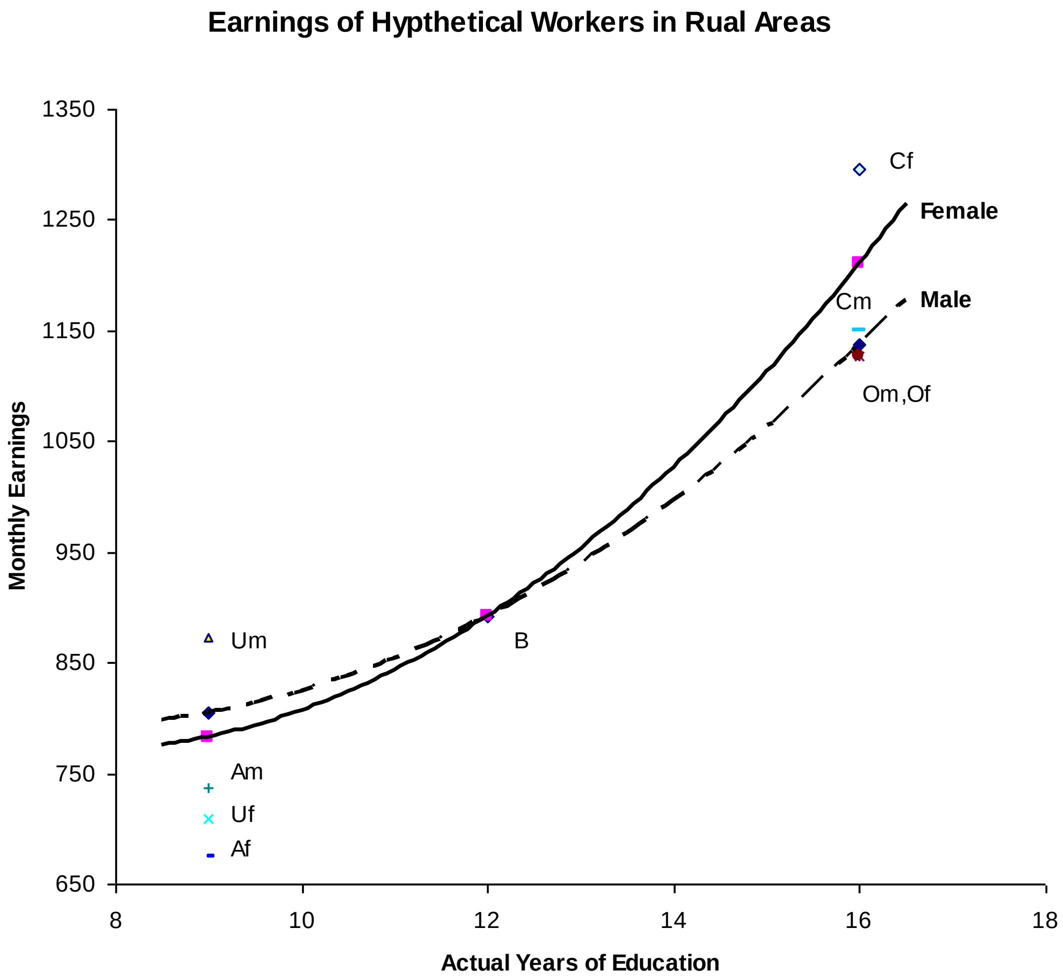


Table 7

Implied Returns to Education, Adjusting for Effects of ORU Variables, Comparisons of Males and Females in Rural Areas

	% Payoff
Males	3.39
Females -no adjustment	7.68
(i) Assuming the same earnings effects to required education, under education and over education categories as for males	3.42
(ii) as for (i) but also assuming the same levels of required education, under education and over education as for males	3.35
(iii) as for (ii) but also assuming the same distribution across education categories for females as for males	3.39

Source: Authors’ calculations.

The implied return to education computed using the predicted mean earnings at each level of education under the Chiswick and Miller (2008) method is 3.39 percent for males, which is a little different from that computed via the Mincerian equation, 3.81 percent. The return to education for females is 7.68 percent, which is also different from that calculated via the Mincerian equation, 7.18 percent. These divergences are due to the relatively small samples, as noted by Chiswick and Miller (2008)(2010a). However, the decomposition can still be used to analyse the reasons for the gender difference in the return to schooling.

The first adjustment considered in Table 7 is for gender differences in the partial effects of required education, over education and under education on earnings. This involves replacing the coefficients of required education (9.35 percent), over education (5.83 percent) and under education (-7.63 percent) for females by the 6.35 percent, 5.78 percent and -0.75 percent values for males. Following this, predicted mean earnings at each level of schooling are used to compute an adjusted payoff to schooling. This adjusted return to education is found to be 3.42 percent, 4.26 percentage points lower than the 7.68 percent payoff without adjustment. In other words, the gender differences in the partial effects on earnings associated with required education, over education and under education appear to be the major reason for the difference between the return to education for males and females.

The next adjustment investigates the effect of gender differences in the distribution of workers across the required education, over education and under education categories at each level of education. As noted in Section III, this involves replacing the means of required education, over education and under education at each level of education of females by those of males. Then, through similar weighted regression, a new implied return to education is determined, and this is 3.35 percent.

This is 0.08 of a percentage point lower than that obtained with only adjustment for the gender differences in the returns to the ORU variables. Thus, the different distributions of males and females across the required education, over education and under education categories at the different years of actual education have only a minor effect on the difference in the return to education between males and females.

Finally, the weighted regression based on the adjusted mean earnings in the previous step is re-estimated using the number of males at each level of education as weights. The implied return to education increases slightly, from 3.35 percent to 3.39 percent, which is, as expected under this methodology, the return to education for males. The 0.04 percentage point increase means that the different distributions across the various levels of actual schooling have almost no effect on the gender difference in the returns to education.

To sum up, 4.26 percentage points of the difference in the return to education for males and females is due to the differences between males and females in the partial effects on earnings of the required education, over education and under education variables. Only 0.08 percentage point is due to the different distributions of male and female workers across the required education, over education and under education categories. The differences in the distributions of males and females across the years of schooling categories employed in the analysis are associated with a 0.04 percentage point effect which works in the opposite direction to the previous two effects.

Therefore, the partial effects on earnings associated with the ORU variables are the major reason for the gap in the return to education between males and females. It is thus useful to investigate the relative importance of the required education, over education and under education variables in this regard. Rather than replacing the three coefficients of required education, under education and over education of females by those of males simultaneously, they are replaced in sequence. Relevant results are presented in Table 8. These show that the adjustment for the return to required education is associated with a narrowing of the gap in the return to actual years of schooling by about 0.81 of a percentage point. Adjustment for the earnings effects of under education is linked to a decrease of the gap in the return to education by 3.44 percentage points. Thus, the earnings effects of under education contribute about 80 percent and those of required education contribute another 19 percent to the higher return to schooling for females. The very minor gender difference in the earnings effects of over education, however, does not contribute to the gender differences in the return to schooling for females.

Table 8

Implied Returns to Education, Detailed Adjustment for Effects of ORU Variables, Comparisons of Males and Females in Rural Areas

	% Payoff
Males	3.39
Females	
-no adjustment	7.68
(i) Assuming the same earnings effects to required education as for males	6.87
(ii) Assuming the same earnings effects to required education and under education as for males	3.43
(iii) Assuming the same earnings effects to required education, under education and over education as for males	3.42
(iv) as for (iii) but also assuming the same levels of required education, under education and over education categories as for males	3.35
(v) as for (iv) but also assuming the same distribution across education categories for females as for males	3.39

Source: Authors' calculations.

(ii) Sensitivity Analysis

The results of the adjustments described above may be sensitive to the order in which they are undertaken. This is termed path dependence (see Chiswick and Miller, 2010a). There are two types of path dependence that may be of some consequence here. First, in relation to Table 7, it may matter whether the adjustments described in step (ii) are undertaken before those described in step (i). Second, in relation to Table 8, it may matter whether the adjustments described in steps (i), (ii) and (iii) are undertaken in a different order. The potential impact of both sources of path dependence was examined by changing the order in which the various adjustments were considered. Thus in the first instance the computations in Table 7 were undertaken by first adjusting the mean values of the ORU variables and then adjusting the coefficients of the ORU model. Relevant results can be found in Appendix Table 2 in Ren and Miller (2010).

The results of this set of calculations also show that the earnings effects of over education, required education and under education are much larger than the effects of the distributions across the over education, required education and under education categories when the order has changed. In this regard the pattern of effects in Ren and Miller (2010) is the same as that evident in Table 7. The change in the order of the first two, however, is associated with an increase of 0.48 percentage point in the value of the first adjustment made and a concomitant fall of the same amount in the value of the second adjustment made.

The second assessment of the importance of path dependence involves changing the order of the first three adjustments in Table 8. In Table 8 the sequence of replacements was the coefficient of required education, that of under education and then that of over education. This order was changed to required education, over education and then under education. The detailed calculations are presented in Appendix Table 3 in Ren and Miller (2010). This test revealed that the result is not dependent on the order in which this sequence of adjustments is made.

(iii) Interpretation

As the gender differences in the effects on earnings of required education and under education result in females having a lower payoff to actual years of schooling, and the gender difference in the effects on earnings of over education are linked to a lessening, albeit inconsequential, of this advantage, it is important to examine why these differences in the ORU model arise. As described in Section I, self selection of females, the comparative disadvantage in unskilled occupations of females and the limited supply of skilled female workers are generally considered as the major reasons for the higher return to education for females in the literature. Table 9 illustrates how these possible reasons for the higher return to education may impact the ORU analysis.

Table 9

ORU-Based Explanations for the Gender Differential in the Payoff to Schooling in China

Explanation for gender difference in return to schooling	Likely effect in labour market	Likely effect in ORU model	Expected impact on females' return to education
Female workforce more highly self-selected	Female workers more able at each level of schooling	Higher payoff to required education and over education for females	Higher payoff to education for females
Limited supply of skilled female workers	Relatively fewer skilled female workers in the labour market	Higher payoff to required education for females	Higher payoff to education for females
Technological requirements of jobs: males have comparative advantage in manual work	Over qualified females disadvantaged and correctly matched females in high-skilled jobs advantaged	Higher payoff to correctly matched education and lower payoff to over education for females	Ambiguous
Selection more favourable at high levels of schooling	Under qualified females disadvantaged	Larger penalty to under education for females	Higher payoff to education for females
Discrimination against under educated females	Under qualified females disadvantaged	Larger penalty to under education for females	Higher payoff to education for females

The first row in Table 9 relates to the self-selection argument. It suggests that favourable self-selection results in more able females at each level of schooling entering the labour market and less able females staying outside the labour market. Ability as an omitted variable in the ORU estimating equation should therefore be associated with a higher return to required education for females than for males (as shown in Table 4). This, in turn, is shown in Table 8 to be associated with a higher payoff to education for females. Moreover, any tendency for this selection mechanism to be more intense at higher levels of schooling will accentuate the omitted variable (ability) bias and hence exaggerate the gender difference in the payoffs to both required and actual years of education.

Second, because of the deeply rooted Confucian concept that ‘boys are better than girls’, the opportunities for higher education have been fewer for females than for males, which reduces the relative supply of highly skilled females. Thus, females with higher education become scarce resources in the labour market, which also causes the return to required education to be higher for females than for males. This again will be associated with a higher payoff to actual years of schooling for males than for females.

Third, females’ comparative disadvantage in physical strength, which is often required in low-skilled occupations with low required levels of education, will be associated with a lower payoff to any surplus schooling for females who gain employment in those occupations. This will be associated with a lower payoff to schooling for females than for males.

Fourth, to the extent that, as hypothesised above, female labour force participants at higher levels of schooling are more intensely selected than those at lower levels of schooling, under educated females workers will have greater difficulty competing with their correctly matched counterparts. This would be associated with a higher earnings penalty for years of under education for females than for males (see Table 4) and hence with a higher payoff to schooling (see Table 8).[13] The same pattern of effects will arise where labour market discrimination is more intense against lower educated females, as argued by Dougherty (2005).

VII. CONCLUSION

A higher return to actual years of education for females than that for males in rural China has been documented by a number of researchers (*e.g.*, Johnson and Chow, 1997; Meng, 1998). The findings in the current study follow those reported in these previous studies.

There are also considerable gender differences in the estimated impacts of the required education and under education variables. The return to required education for females in rural China is 9.35 percent, and this is about 3 percentage points higher than that for males, 6.35 percent. In other words, the payoff to correctly matched education in the Chinese rural labour market is much higher for females than it is for males. The return to under education for females is -7.63 percent, which is a more substantial impact than the -0.75 percent impact among males. That is, the wage penalty where workers are under educated in their occupation is quite pronounced for females, but it is relatively minor for males. Both of these factors are shown in the Chiswick and Miller (2008) decomposition to be associated with a higher payoff to education for females than for males in rural areas. The return to over education for males is 5.78 percent, and this is only slightly lower than the 5.83 percent return for females. This differential tends to reduce the payoff to education for females, though the impact is essentially inconsequential. The distributions of correctly matched, over educated and under educated male and female workers were also been considered using the Chiswick and Miller (2008) decomposition, but the results show that relatively little of the gender difference in the return to education could be attributed to this source.

The reasons advanced in previous literature for the gender difference in the payoff to schooling in rural China were linked to the gender difference in the returns to correctly matched education, under education and over education. A highly self-selected female workforce and a limited supply of skilled female workers were associated with the higher return to required education for females. Intense selection at higher levels of schooling means that under qualified females will have greater difficulty competing with their correctly matched counterparts.

ENDNOTES

[1]The literature that compares the return to education in urban and rural areas, but which does not make a distinction between males and females (*e.g.*, Li and Li, 1994; Li C., 2003), will not be considered here.

[2] CHIP was a joint research project between the Institute of Economics, Chinese Academy of Social Sciences and the Ford Foundation. Support was also provided by the Columbia East Asian Institute and the City University of New York. The sample includes observations from rural areas in all of the 28 provinces other than Tibet, Xinjiang and Taiwan. Three surveys of this project were conducted in 1988, 1995 and 2001.

[3] In comparison, in urban areas the return to education was 4.46 percent for females and 2.78 percent for males.

[4] Four counties were involved in this survey: Wuxi county in Jiangsu province, Nanhai county in Guangdong province, Jieshou county in Anhui province and Shangrao county in Jiangxi province.

[5] This measure follows Chen and Hamori (2009), who use the same data set, but is narrower than the measure of earnings in Zhang *et al.* (2005) and Li H. (2003), where bonuses, subsidies and other labour-related income were included along with basic monthly earnings.

[6] The recent analyses by Chiswick and Miller (2010b) indicate that the findings in this literature are quite robust to the choice of reference years of education.

[7] This earnings effect is computed using the algorithm proposed by Halvorsen and Palmquist (1980), namely g=exp(c)-1, where g represents the percentage effect on the dependent variable and c represents the coefficient of a dummy variable.

[8] Treating individuals who do not report valid wage data in this study of selection bias as non-participants suggests that these data are missing at random.

[9] As ln(892)=6.7935, these figures are computed as earnings = exp(6.7935-3*0.06348) and earnings =exp(6.7935+4*0. 0.06348), respectively.

[10]earnings =exp(6.7935-3*0.09346) and earnings =exp(6.7935+4*0.09346), respectively.

[11]earnings =exp(6.7935-3*0.00749) and earnings =exp(6.7935-3*0.07633), respectively.

[12] earnings =exp(6.7935+4*0.05778) and earnings =exp(6.7935+4*0.05832), respectively.

[13] The earnings effects associated with correctly matched education and under education for males (0.0635 and -0.0075) and for females (0.0935 and -0.0763) mean that the earnings positions of under educated male and female workers are similar (they differ by around one percentage point per year of schooling). Hence, the greater earnings penalty associated with under education among females can be viewed simply as a consequence of their failure to reap the greater rewards for matching on the basis of level of schooling in the female labour market.

REFERENCES

Chen, G. and Hamori, S. (2009) Economic returns to schooling in urban China: OLS and the instrumental variables approach. *China Economic Review*, 20, pp. 143-152.

Chiswick, B. R. and Miller, P. W. (2008) Why is the payoff to schooling smaller for immigrants?. *Labor Economics*, 15, pp. 1317-1340.

Chiswick, B. R. and Miller, P. W. (2010a) An explanation for the lower payoff to schooling for immigrants in the Canadian labour market. pp.41-76 in T. McDonald, E. Ruddick, A. Sweetman and C. Worswick (eds.), *Canadian Immigration: Economic Evidence for a Dynamic Policy Environment* (Montreal: McGill-Queen's University Press).

Chiswick, B. R. and Miller, P. W. (2010b) Does the choice of reference levels of education matter in the ORU earnings equation?. *Economics of Education Review* (forthcoming).

Choudhury, S. (1994) New evidence on public sector wage differentials. *Applied Economics*, 26, pp. 259-266.

Dougherty, C. (2005) Why are the returns to schooling higher for women than for men?. *Journal of Human Resources*, 40, pp. 969-988.

Deolalikar, A. B. (1993) Gender differences in the returns to schooling and school enrollment rates in Indonesia. *Journal of Human Resources*, 28, pp. 899-932.

Gyourko, J. and Tracy, J. (1988) An analysis of public- and private-sector wages allowing for endogenous choices of both government and union status. *Journal of Labor Economics*, 6, pp. 229-253.

Halvorsen, R. and Palmquist R. (1980) The interpretation of dummy variables in semilogarithmic equations. *American Economic Review*, 70, pp. 474-475.

Hartog, J. (2000) Over-education and earnings: where are we, where should we go?. *Economics of Education Review*, 19, pp. 131-147.

Hartog, J. and Oosterbeek, H. (1993) Public and private sector wages in the Netherlands. *European Economic Review*, 37, pp. 97-114.

Hou, F. (2004) The study on rates of investment return to human capital in China rural areas. *Economics Research Journal*, 12, pp. 75-84.

Johnson, E. N. and Chow, G. C. (1997) Rates of return to schooling in China. *Pacific Economic Review*, 2, pp. 101-113.

Lee, L. (1983) Generalized econometric models with selectivity. *Econometrica*, 51, pp. 507-512.

Li, C. (2003) How does culture level influence the income of people-the study on current return to education. *Sociology Studies*, 3, pp. 64-76.

Li, H. (2003) Economic transition and returns to education in China. *Economics of Education Review,* 22, pp. 317-328.

Li, S. and Li, W. (1994) The estimation of individual returns to education in China's education investment". pp. 442-456 in R. Zhao, S. Li and C. Riskin (eds), *Studies on Income Distribution of Chinese Residents,* (China: China Society Science Press).

Meng, X. (1998). Male-female wage determination and gender wage discrimination in China's rural industrial sector. *Labour Economics*, 5, pp. 67-89.

Miller, P.W. (1987) Gender differences in offered and observed wages in Canada, 1980. *Canadian Journal of Economics*, 20, pp.225-244.

Miller, P.W. (2009) The gender pay gap in the US: does sector make a difference?. *Journal of Labor Research*, 30, pp.52-74.

Puhani, P.A. (2000) The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14, pp. 53-68.

Ren, W. and Miller, P.W. (2010) Gender differentials in the payoff to schooling in China. Institute for the Study of Labor Discussion Paper No. 5179 (Bonn: Institute for the Study of Labour).

Sun, Z. (2002) The empirical analysis in family education decision in China rural areas—the study based on Gansu survey. PhD dissertation, *Peking Normal University*.

Yao, X. and Zhang H. (2004) An analysis of the differential of the rates of return to schooling between urban and rural labor in China-evidence from the survey in Zhejiang, Guangdong, Hunan and Anhui. *Collected Essays on Finance and Economics*, 6, pp. 1-7.

Zhang, J., Zhao Y., Park A. and Song, X. (2005) Economic returns to schooling in urban China, 1988 to 2001. *Journal of Comparative Economics,* 33, pp. 730-752.

Zhao, Y. (1997) Labor migration and returns to rural education in China. *American Journal of Agricultural Economics*, 79, pp. 1278-1287.

APPENDIX A

Means and Standard Deviations of Variables by Gender

Variable	Males	Females
Log Earnings	6.839 (0.594)	6.582 (0.603)
Years of Education	10.143 (2.866)	9.879 (3.292)
Required education	10.148 (1.479)	10.134 (1.474)
Over education	0.852 (1.440)	0.718 (1.284)
Under education	0.857 (1.641)	0.973 (2.033)
Log hours	3.820 (0.327)	3.817 (0.405)
Experience	25.067 (10.843)	21.436 (10.287)
Married	0.893 (0.310)	0.866 (0.341)
Ruralreg	0.453 (0.498)	0.411 (0.493)
Central(1)	0.321 (0.467)	0.268 (0.444)
East	0.476 (0.500)	0.518 (0.500)
Stat(2)	0.403 (0.491)	0.351 (0.478)
Prov	0.450 (0.498)	0.499 (0.501)
Clerk(3)	0.104 (0.306)	0.140 (0.347)
Junior	0.274 (0.447)	0.274 0.447)
Senior	0.103 (0.304)	0.049 (0.217)
Leader	0.106 (0.308)	0.031 (0.171)
Other	0.046 (0.209)	0.022 (0.147)

Notes: (1) West is the benchmark region; (2) Collective enterprise is the benchmark ownership; (3) The benchmark occupation is unskilled workers (ordinary labourers, service workers, farm workers).