

IZA DP No. 5179

## Gender Differentials in the Payoff to Schooling in China

Weiwei Ren  
Paul W. Miller

September 2010

# Gender Differentials in the Payoff to Schooling in China

**Weiwei Ren**

*University of Western Australia*

**Paul W. Miller**

*Curtin University  
and IZA*

Discussion Paper No. 5179  
September 2010

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Gender Differentials in the Payoff to Schooling in China<sup>\*</sup>

This paper examines the gender differential in the payoff to schooling in China. The analyses are conducted separately for rural and urban areas, and 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. Both of these factors are shown to be linked to the higher payoff to schooling for females than for males. Over educated females, however, are advantaged compared with their male counterparts, though this has little effect on the differential in the payoff to schooling between males and females in rural China. 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. The payoffs to actual years of schooling for males and females in urban China are remarkably similar in this study.

JEL Classification: J31, J62, J70

Keywords: China, schooling, earnings, rates of return

Corresponding author:

Paul W. Miller  
School of Economics and Finance  
Curtin University  
GPO Box U1987  
Perth, WA 6845  
Australia  
E-mail: [Paul.Miller@curtin.edu.au](mailto:Paul.Miller@curtin.edu.au)

---

<sup>\*</sup> 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.

## **GENDER DIFFERENTIALS IN THE PAYOFF TO SCHOOLING IN CHINA**

### **I. INTRODUCTION**

Two empirical regularities in studies of the determinants of earnings in China are that the return to education in urban areas exceeds that in rural areas, and the return 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 higher than in Meng's (1998) study for rural areas, but there was again clear evidence of education being associated with a higher payoff for females (6.9 percent) than for males (4.3 percent).

This gender differential in the payoff to schooling carries across to disaggregated analyses. Wang and Cai (2008), for example, based their analysis around the 16 sectors in the Chinese economy identified in the China Statistical Year Book 2002. They used mean wages to combine these 16 sectors into four reasonably homogeneous groups, and analyzed the return to education in each of these groups. In three of the four sector groups the payoff to education was higher for females than for males (by between one and six percentage points). In the remaining group the return to education for males was one percentage point higher than that for females.

Moreover, the gender differential in the payoff to schooling in China appears to have widened over time. This pattern has been reported by both Yu (2000) and Zhang *et al.* (2005). Zhang *et al.* (2005), for example, reported that the gender gap in the returns to education was 2.3 percent in 1988, 1.9 percent in 1992 and 4.8 percent in 2001.

Several explanations have been proposed for the higher payoff to schooling for females in the Chinese labor market. 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 labor force relative to men, whose labor force participation is nearly universal. This more intense positive selection into labor market activity among women would result in the ability of labor market entrants being higher than that of women outside the labor 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 to investigate the reasons for the differential in the payoff to education between males and females in China. This literature proposes that there is a usual or reference level of education for each occupation.<sup>1</sup> The over

---

<sup>1</sup> Workers who have this level of education are viewed as being correctly matched to the educational requirements of their jobs. Workers with fewer years of education than that which is usual in their occupation are viewed as being under educated. Workers with more years of education than that which is usual in their occupation are viewed as being over educated.

education/under education literature has shown that the earnings of workers will vary according to whether they are correctly matched, over educated or 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 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, VI and VII present the empirical results. The first two of these sections are for rural areas and the final one is for urban areas. For rural areas the first section 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. The second section for rural areas 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. A parallel set of analyses for urban areas is presented in Section VII. Section VIII 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, 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.

The most widely-used data for this type of study in urban areas are the first two waves of the Chinese Household Income Project (CHIP), conducted in 1988 and 1995.<sup>2</sup> Li and Li (1994) used the first of these waves. Based on a simple specification of the earnings equation that contained only schooling, experience and experience squared, they reported that the return to education in urban areas was 2.5 percent for males and 3.7 percent for females.

Lai (1998) extended this research by utilizing the 1995 wave of CHIP. Using the same model as Li and Li (1994), Lai (1998) reported that the return to education for males had increased to 5.14 percent by 1995 and the return to education for females had increased to 5.99 percent, though females still maintained an advantage in the payoff to education.

A similar pattern emerges from Li H.'s (2003) augmented earnings function that included dummy variables for ownership type, industry, province, Party

---

<sup>2</sup> 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. The sample's urban observations were obtained from Liaoning, Shanxi, Jiangsu, Guangdong, Anhui, Hunan, Hebei, Gansu and Yunnan provinces as well as from Beijing. Three surveys of this project were conducted in 1988, 1995 and 2001.

membership, ethnic minority and for youth “re-education”. This was applied to a sample of urban households from the State Statistical Bureau’s panel in CHIP (1995), and the return to education was reported to be 6.9 percent for females and 4.3 percent for males.

A higher payoff to schooling for females than for males in urban areas has also been reported in analyses of other data sets. Using the China Urban Households Survey (CUHS) conducted by China’s National Bureau of Statistics from 1988 through 2001,<sup>3</sup> Zhang *et al.* (2005) reported that the return to schooling for females exceeded that for males by an average of about 60 percent over this thirteen-year period, and the gender differential in payoffs has tended to widen over time. Thus, while the return to education for females increased from 5.2 percent in 1988 to 13.2 percent in 2001, that for males only increased from 2.9 percent in 1988 to 8.4 percent in 2001. Similar changes over time are presented in Chen and Ju’s (2004) analysis of these data for the period 1996-2000, and in Li and Ding’s (2003) analysis of the 1995 and 1999 surveys conducted by the Income Distribution Group and Urban Poverty Group of the Institute of Economic Studies at the Chinese Academy of Social Sciences.<sup>4</sup>

---

<sup>3</sup> The CUHS is conducted by the Urban Socio-Economics Survey Organization of the State Statistical Bureau. It covers 103 cities and 80 counties in two municipalities (Beijing and Chongqing) and 8 provinces (Shaanxi, Liaoning, Jiangsu, Anhui, Hubei, Guangdong, Sichuan and Gansu).

<sup>4</sup> The first sample survey covers the municipality of Beijing and ten provinces, namely Shanxi and Gansu in Western areas, Liaoning in Northwest areas, Jiangsu and Guangdong in coastal areas, Anhui, Hubei and Henan in Mid areas and Sichuan and Yunnan in Southwest areas. It surveys the households from 1990-1995. The second sample is drawn from the first one and covers Beijing and five provinces, which are Gansu in the West, Liaoning in the Northeast, Jiangsu in the coastal regions, Henan in the Central areas and Sichuan in the Southwest. The second survey covers 1996-1999.

Wang and Cai's (2008) study of gender wage differentials in China's urban labor market was based on data from the China Urban Labor Survey 2001 (CULS).<sup>5</sup> They based their analysis on the 16 sectors in the Chinese economy identified in the China Statistical Year Book 2002. The authors combined these 16 sectors into four groups that had similar mean wages. In three of these four groups the payoff to schooling for females exceeded that for males (by between one and six percentage points). In the remaining group, the payoff to schooling for males (12.5 percent) exceeded that for females (11.2 percent).

A higher payoff to schooling for females than for males has not, however, been reported for all studies of earnings in urban China. Hence in a recent study, based on data from the Chinese Health and Nutrition Survey (CHNS) for 2004 and 2006, Chen and Hamori (2009) found the return to education for males (at 8.06 percent) was very similar to that for females (7.67 percent).<sup>6</sup>

The CHIP 1988 data set has also been used to examine the return to education in rural areas.<sup>7</sup> Thus Johnson and Chow (1997) calculated the return to education using a pooled sample of urban and rural workers and separately for urban workers and rural workers. They reported that in urban areas the return to education was 4.46 percent for females and 2.78 percent for males. In rural areas the returns to education for females and males were 4.82 percent and 2.95 percent, respectively.

---

<sup>5</sup> This survey was conducted by the Institute of Population and Labor Economics at the Chinese Academy of Social Sciences and covered five cities: Fuzhou, Shanghai, Shenyang, Wuhan and Xi'an. The sample is smaller than CHIP and CUHS.

<sup>6</sup> The data set they used is the same as the data set used in the current paper. Further details can be found in Section IV.

<sup>7</sup> Some literature (Li and Li, 1994; Li C., 2003) has considered the return to education in rural areas and that in urban areas separately. However, these studies generally do not compare the return to education between males and females in rural areas but rather focus on differences in the return to education between urban area and rural area. For example, Li and Li (1994) found the sector gap in the return to education is almost two percentage points and Li C. (2003) found the gap is seven percentage points.

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.<sup>8</sup> 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 second 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, and this was less than the return to education for males (of 3.9 percent) in rural areas.

Hence, while most studies report that the payoff to schooling for females in China typically exceeds that for males, there are a number of studies that report the opposite finding. The reasons for these gender differentials in the payoff to schooling in China 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 (5.2 percent) than for the

---

<sup>8</sup> Four counties were involved in this survey: Wuxi county in Jiangsu province, Nanhai county in Guangdong province, Jiesshou county in Anhui province and Shangrao county in Jiangxi province.

native born (10.6 percent) in the US. The starting point for their study is the ORU (Over education/Required education/Under education) model of earnings determination. In this model 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, labor market experience (EXP) and its square, and other variables, such as location and marital status, that are usually held to affect earnings. That is:

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

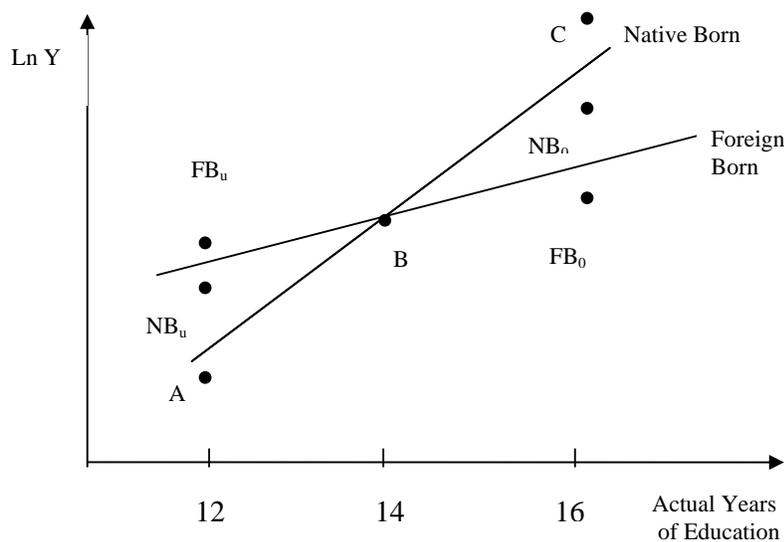
This model was estimated separately by Chiswick and Miller (2008) for samples of the native born and foreign born. The authors reported that the return on the reference years of education was 15.4 percent for the native born, and an almost identical amount, 15.3 percent, for the foreign born. A year of over education for the native born was associated with 5.6 percent higher earnings and that for the foreign born was associated with an increase in earnings of only 4.4 percent. Years of under education were associated with an earnings penalty of -6.7 percent among the native born and an earnings penalty of only -2.1 percent among the foreign born.

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 have smaller gains associated with ‘surplus’ education than the native born, which is represented as  $FB_0$  and  $NB_0$  in this 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 have bigger gains associated with ‘deficit’ education than the native born, which is represented as  $FB_u$  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.

This discussion above 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. This is the information that Chiswick and Miller (2008) used in their decomposition of the difference in the payoff to education for the native born and foreign born. Their method is 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. As described above this approach mimics the usual calculation of the Mincerian returns to education. The return to schooling for the foreign born computed under this exercise was 5.3 percent.

A comparable approach for the native born yielded a return to schooling of 10.5 percent.

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 the mean predicted earnings at each level of education on the years of education was again estimated. Under this scenario the payoff to schooling was 8.5 percent, which is 3.2 percentage points more than the Mincerian return to schooling for the foreign born. This result illustrates that about two thirds of the difference in the payoffs to schooling for the foreign born and native born is due to the differences between these birthplace groups in the partial effects on earnings associated with over education and under education.

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. This payoff to schooling for the foreign born was 8.6 percent, which is 0.1 of a percentage point higher than that for the second prediction. This result means that only three percent of the difference in the payoffs to education for the native-born and foreign-born workers is due to different distributions of workers across the over education/under education categories.

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. As expected, this regression yields the payoff to schooling for the native born. This informs that 36 percent of the difference in the payoff to education for the two groups in the Chiswick and Miller (2008) study is due to 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 currently select a sample of about 4400 households with a total of 19000 individuals.

Because of the focus on the returns to education in this paper, only individuals 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. However, as the literature reviewed in Section II has documented differences in the determinants of earnings between rural and urban areas, separate analyses are undertaken for workers in these regions. The potential for this partition of the sample 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 labor 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.<sup>9</sup>

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, 1919 observations remain, of which 979 are for rural workers, including 614 males and 365 females, and 940 for urban workers, including 536 males and 404 females. All the analyses below are based on these observations.

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 utilize their levels of education in their work; and Realized 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.<sup>10</sup>

---

<sup>9</sup> 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 labor-related income were included along with basic monthly earnings.

<sup>10</sup> 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.

## V. EMPIRICAL RESULTS FOR RURAL AREAS

The processes of wage determination in rural areas and urban areas in China are very different, and because of this we consider the empirical results and decomposition in these areas separately.<sup>11</sup> We focus on rural areas in some detail in Sections V and VI, and only briefly discuss the main issues for urban areas in Section VII. 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.

### (i) Mincerian equation results

Table 1 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).

---

<sup>11</sup> As noted in the previous section, possible links between this partition of the sample and sample selection bias are considered as part of the sensitivity analyses later in this section.

**Table 1**  
**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 <sup>2</sup> /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 <sup>2</sup>	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.<sup>12</sup> 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

---

<sup>12</sup> 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.

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

Utilizing the Realized Matches method, Table 2 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
<i>Sample size</i>	<i>979</i>	<i>614</i>	<i>365</i>

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

There are two main features of Table 2. 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. Secondly, 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 3 lists estimates of the ORU model that explore this issue.

Columns (i) and (ii) of Table 3 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 3, 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 3 can focus on the estimated effects associated with the three ORU variables.

**Table 3**  
**Estimates of ORU Model of Earnings Determination in Rural Areas**

Variables	All workers		Males	Females
	(i)	(ii)	(iii)	(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 <sup>2</sup> /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 3, 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 1, 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 3 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 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 labor 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 labor 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 labor market experience variables were also considered, and the findings are noted below). However, the sample selection term was significant in only one equation. Estimates of the lambda term and its 't' statistic are presented in Table 4. For reference purposes, this table also contains results for the urban area. The full set of results is available from the authors upon request.

It is apparent from Table 4 that there is limited evidence in favor 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.

**Table 4**  
**Estimates of Coefficient on Sample Selection Term, Various Models**

Sample and Model	All Workers		Males	Females
	(i)	(ii)	(iii)	(iv)
<b>Rural Areas</b>				
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)
<b>Urban Areas</b>				
Mincer	0.0169 (0.05)	-0.0409 (0.13)	0.4059 (0.68)	-0.0467 (0.13)
ORU	-0.0071 (0.04)	0.0657 (0.35)	0.1676 (0.60)	-0.1209 (0.48)

*Notes:* 't' statistics are presented in parentheses. \*\* = significant at the 5 percent level.

*Source:* Statistical Appendix, available upon request.

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 5. 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.

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.<sup>13</sup> The monthly earnings of Type A female workers and Type C female workers are predicted as 674 yuan and 1296 yuan, respectively.<sup>14</sup>

**Table 5**  
**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

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.<sup>15</sup> Similarly, we can predict the monthly earnings of over educated, Type O, workers as 1124 yuan for males and 1126 yuan for females.<sup>16</sup>

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 figure reveals the distinctive patterns from the ORU literature and illustrates how males and females differ in key ways in the earnings effects associated with mismatched education.

<sup>13</sup> As  $\ln(892)=6.7935$ , these figures are computed as earnings =  $\exp(6.7935-3*0.06348)$  and earnings =  $\exp(6.7935+4*0.06348)$ , respectively.

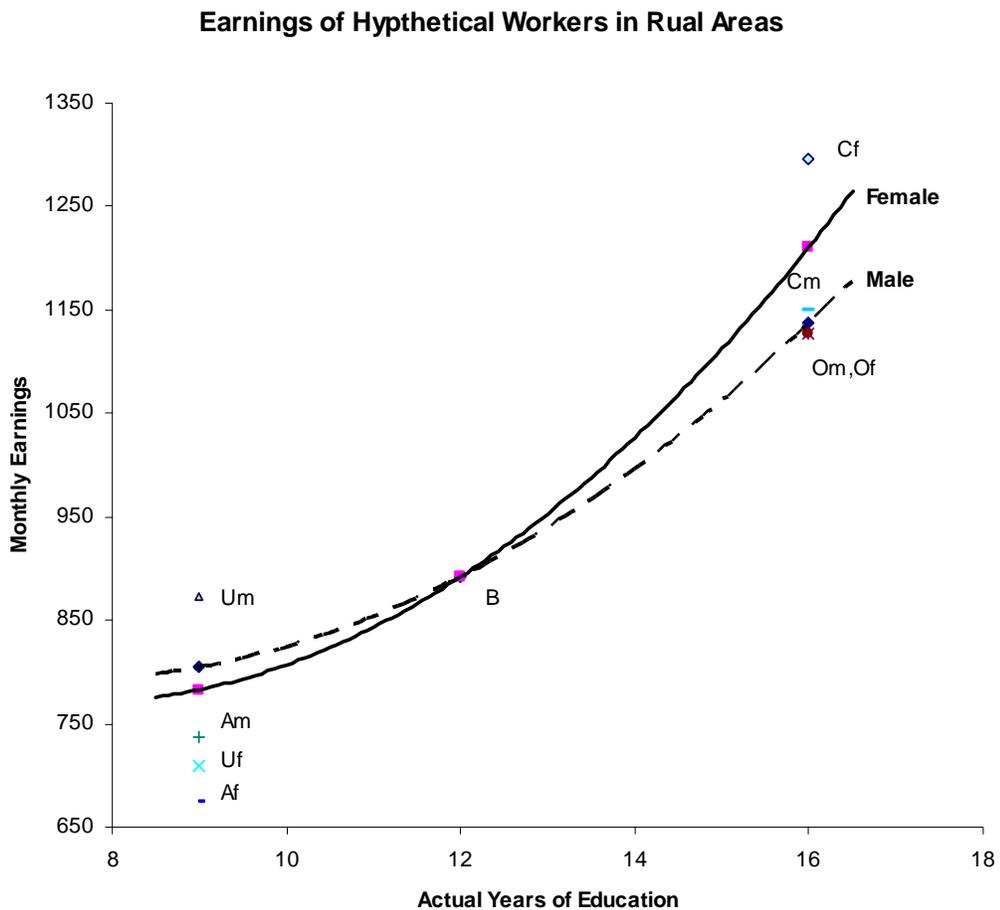
<sup>14</sup> earnings =  $\exp(6.7935-3*0.09346)$  and earnings =  $\exp(6.7935+4*0.09346)$ , respectively.

<sup>15</sup> earnings =  $\exp(6.7935-3*0.00749)$  and earnings =  $\exp(6.7935-3*0.07633)$ , respectively.

<sup>16</sup> earnings =  $\exp(6.7935+4*0.05778)$  and earnings =  $\exp(6.7935+4*0.05832)$ , respectively.

First, there are considerable earnings increments to correctly matched education, as shown by a comparison of workers of Types A, B and C. Second, the earnings of the over-educated Type O workers are lower than the earnings of Type C individuals who are correctly matched in jobs that require exactly 16 years of education, but the differences are small. This small gap in the earnings of Type C and Type O workers contrasts with findings in other studies (Hartog, 2000; Chiswick and Miller, 2008) where the over educated are much more disadvantaged. Moreover, there is a very minor difference in terms of earnings between Type O males and females.

**Figure 2**



Third, the Type U workers, with 9 years of education, but working in an occupation that requires 12 years of education, earn more than the Type A workers, with 9 years of education and working in an occupation that requires 9 years of education. However, they earn less than Type B workers who are correctly matched in occupations that require 12 years of education. Comparing Type U and Type A workers, it is seen that under educated female workers have earnings five percent above the earnings of correctly matched female workers with 9 years of schooling. However, the respective earnings effect among males is an 18 percent earnings advantage.

## VI. DECOMPOSITION ANALYSIS

### (i) Basic Results

In this section, the reasons for the differential in the return to education between males and females in China's rural areas are investigated using the decomposition proposed by Chiswick and Miller (2008), which was described in Section III. Results from application of this decomposition to the results from the estimation of the ORU earnings model for rural China are given in Table 6.

**Table 6**  
**Implied Returns to Education, Adjusting for Effects of ORU Variables,**  
**Comparisons of Males and Females in Rural Areas**

	% Payoff
Males	3.39
Females	7.68
-no adjustment	
(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 still can be used to analyze the reasons for the gender difference in the return to schooling.

The first adjustment considered in Table 6 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 7. 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 7**  
**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 6, it may matter whether the adjustments described in step (ii) are undertaken before those described in step (i). Second, in relation to Table 7, 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 6 were undertaken by first adjusting the mean values of the ORU variables and then adjusting the coefficients of the ORU model. Relevant results are reported in Appendix Table 2.

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 Appendix Table 2 is the same as that evident in Table 6. 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 7. In Table 7 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. Detailed calculations are presented in Appendix Table 3. 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 it 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 8 illustrates how these possible reasons for the higher return to education may impact the ORU analysis.

**Table 8**

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

<b>Explanation for gender difference in return to schooling</b>	<b>Likely effect in labor market</b>	<b>Likely effect in ORU model</b>	<b>Expected impact on females' return to education</b>
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 labor 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 favorable 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 8 relates to the self-selection argument. It suggests that favorable self-selection results in more able females at each level of schooling entering the labor market and less able females staying outside the labor 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 3). 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 labor 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 hypothesized above, female labor 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 3) and hence with a higher payoff to schooling (see Table 7).<sup>17</sup> The same pattern of effects will arise where labor market discrimination is more intense against lower educated females, as argued by Dougherty (2005).

## **VII. EMPIRICAL RESULTS AND DECOMPOSITION FOR URBAN AREAS**

### **(i) Mincer equation results**

Table 9 presents the estimated coefficients obtained from applying a standard schooling and experience model to urban areas. The payoff to schooling in urban areas is 7.74 percent, which is 2.5 percentage points higher than that in rural areas. This differential is consistent with previous literature. However, unlike most literature in this area, and in contrast to the situation in rural areas, the payoff to schooling for males in urban areas, at 7.50 percent, is very similar to that for females, 7.44 percent. Only the recent study by Chen and Harmori (2009) reports that men and women receive similar payoffs to schooling in China's urban areas.

---

<sup>17</sup> 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 labor market.

**Table 9**  
**Estimates of Mincerian Model of Earnings Determination in Urban Areas**

Variables	All Workers		Males	Females
	(i)	(ii)	(iii)	(iv)
Constant	5.5395*** (20.23)	5.6146*** (21.27)	5.5907*** (16.86)	5.3469*** (10.57)
Years of education	0.0774*** (11.33)	0.0567*** (7.74)	0.0750*** (8.06)	0.0744*** (7.58)
Log hours	-0.0397 (0.59)	-0.0221 (0.34)	-0.0596 (0.76)	-0.0045 (0.04)
Experience (Exp)	0.0165** (2.10)	0.0157* (1.94)	0.0197* (1.76)	0.0097 (0.87)
Exp <sup>2</sup> /100	-0.0154 (1.06)	-0.0201 (1.36)	-0.2427 (1.13)	-0.0004 (0.02)
Female	-0.2252*** (6.29)	-0.2015*** (5.89)		
Married	-0.0519 (0.65)	-0.0715 (0.90)	0.0404 (0.45)	-0.1484 (1.18)
Ruralreg	0.0641 (1.07)	-0.0625 (1.02)	0.0152 (0.18)	0.1122 (1.36)
Central	0.2063*** (3.87)	0.1989*** (3.92)	0.2139*** (2.79)	0.1794*** (2.70)
East	0.2489*** (4.16)	0.2435*** (4.19)	0.2427*** (2.99)	0.2245*** (2.58)
Stat	0.3487*** (6.88)	0.3270*** (6.26)	0.2774*** (3.84)	0.4270*** (6.04)
Prov	0.2895*** (5.80)	0.3245*** (6.31)	0.3004*** (4.21)	0.2529*** (3.52)
Clerk		0.1683*** (2.70)		
Senior		0.3539*** (5.60)		
Junior		0.1426*** (3.09)		
Leader		0.3474*** (5.17)		
Other		0.1162 (1.01)		
Adj. R <sup>2</sup>	0.2260	0.2519	0.1697	0.2579
Sample size	940	940	536	404

*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.

(ii) Returns to Over education, Required education and Under education

From Table 10, it is seen that the wage penalty to under education for males is -4.41 percent, and this differs by two percentage points from the wage penalty for females, -2.41 percent. Females in urban areas have a lower return (6.45 percent) on surplus years of schooling than that for males (7.83 percent). However, the payoff to correctly matched years of schooling for females, 15.07 percent, is around 3.6 percentage points higher than the payoff to correctly matched years of schoolings for males (11.4 percent). The different payoffs to over education and under education for females will tend to lower their payoff to actual years of schooling, while the higher payoff to reference years of schooling will tend to raise the return to actual years of schoolings for females. The fact that the payoffs to actual years of schooling for males and females are similar (see Table 9) suggests that these offsetting factors are of approximately equal magnitudes. This is investigated in the next sub-section.

(iii) Decomposition analysis

Table 11 lists the results from applying the decomposition of Chiswick and Miller (2008) to the minor gender differential in the payoff to schooling for males and females in urban areas. From this table, it is really apparent that each of the adjustment has an inconsequential impact on the implied payoff to schooling for females. These results show that the Chiswick and Miller (2008) decomposition has little explanatory power when there are only minor differences in the return to education for the two groups under consideration (see also Chiswick and Miller (2010a) on this matter).

**Table 10**  
**Estimates of ORU Model of Earnings Determination in Urban Areas**

Variables	All workers		Males	Females
	(i)	(ii)	(iii)	(iv)
Constant	4.8609*** (16.74)	5.2099*** (15.07)	5.1246*** (14.87)	4.2213*** (7.55)
Reference education	0.1311*** (10.37)	0.0907*** (4.02)	0.1143*** (6.69)	0.1507*** (7.71)
Over education	0.0749*** (6.28)	0.0694*** (5.96)	0.0783*** (4.68)	0.0645*** (3.87)
Under education	-0.0372*** (2.84)	-0.0310** (2.34)	-0.0410** (2.43)	-0.0241 (1.22)
Log hours	-0.0155 (0.24)	-0.0112 (0.17)	-0.0495 (0.65)	0.0638 (0.52)
Experience (Exp)	0.0229*** (2.96)	0.0203*** (2.56)	0.0239** (2.11)	0.0203* (1.93)
Exp <sup>2</sup> /100	-0.0333** (2.27)	-0.0314** (2.15)	-0.0363* (1.67)	-0.0294 (1.49)
Female	-0.2398*** (6.74)	-0.2147*** (6.04)		
Married	-0.0752 (0.96)	-0.0761 (0.96)	0.0225 (0.26)	-0.1826 (1.46)
Ruralreg	0.0489 (0.80)	0.0471 (0.76)	0.0002 (0)	0.0938 (1.09)
Central	0.2092*** (4.04)	0.2000*** (3.95)	0.2056*** (2.75)	0.2039*** (3.17)
East	0.2457*** (4.17)	0.2398*** (4.12)	0.2325*** (2.93)	0.2344*** (2.73)
Stat	0.3211*** (6.44)	0.3145*** (6.00)	0.2549*** (3.55)	0.3936*** (5.73)
Prov	0.3283*** (6.61)	0.3304*** (6.48)	0.3353*** (4.67)	0.2971*** (4.19)
Clerk		0.0750 (0.86)		
Senior		0.2661*** (3.20)		
Junior		0.1164** (2.43)		
Leader		0.2432*** (2.63)		
Other		0.1321 (1.16)		
Adj. $R^2$	0.2471	0.2553	0.1804	0.2806
Sample size	940	940	536	404

*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.

**Table 11**  
**Implied Returns to Education, Adjusting for Effects of ORU Variables,**  
**Comparisons of Males and Females in Urban Areas**

	% Payoff
Males	7.41
Females	7.45
-no adjustment	
(i) Assuming the same earnings effects to required education, under education and over education categories as for males	7.54
(ii) as for (i) but also assuming the same levels of required education, under education and over education as for males	7.49
(iii) as for (ii) but also assuming the same distribution across education categories for females as for males	7.41

*Source:* Authors' calculations.

## VIII. CONCLUSION

A higher return to education for females than that for males in China has been documented by many researchers (*e.g.*, Li H. 2003; Wang and Cai 2008). The recent study by Chen and Hamori (2009), however, reported that the return to education for males is 8.06 percent, a little higher than that for females, 7.67 percent. The findings in the current study for rural areas follow those reported in the majority of previous studies, while the finding for urban areas is in accord with the results in Chen and Hamori (2009).

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 labor 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.

The return to education for males is 7.50 percent in urban areas, and this payoff is very similar to that for females, 7.44. The Chiswick and Miller (2008) decomposition was shown to have little explanatory power in this situation. Whether this minor gender difference, which has also been reported by Chen and Hamori (2009), is unique to recent waves or the China Health and Nutrition Survey, or holds for other recently collected data, is a direction for future research.

## REFERENCES

- Chen, G. and Hamori, S., (2009). "Economic Returns to Schooling in Urban China: OLS and the Instrumental Variables Approach", *China Economic Review*, Vol. 20(2), pp. 143-152.
- Chen, L. and Ju, G., (2004). "An Empirical Analysis of Gender Differences in Mincerian Rate of Returns to Education", *Peking University Education Review*, Vol. 2(3), pp. 44-45.
- Chiswick, B. R. and Miller, P. W., (2008). "Why is the Payoff to Schooling Smaller for Immigrants?", *Labor Economics*, Vol. 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", in Ted McDonald, Elizabeth Ruddick, Arthur Sweetman, and Christopher Worswick (eds.), *Canadian Immigration: Economic Evidence for a Dynamic Policy Environment*, Montreal: McGill-Queen's University Press, pp.41-76.
- 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*, Vol. 26, pp. 259-266.
- Dougherty, C., (2005). "Why Are the Returns to Schooling Higher for Women than for Men?", *Journal of Human Resources*, Vol. 40(4), pp. 969-988.
- Deolalikar, A. B., (1993). "Gender Differences in the Returns to Schooling and School Enrollment Rates in Indonesia", *Journal of Human Resources*, Vol. 28(4), 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*, Vol. 6, No. 2, pp. 229-253.
- Halvorsen, R. and Palmquist R., (1980). "The Interpretation of Dummy Variables in Semilogarithmic Equations", *American Economic Review*, Vol. 70(3), pp. 474-475.
- Hartog, J., (2000). "Over-education and Earnings: Where are We, Where Should We Go?", *Economics of Education Review*, Vol. 19(2), pp. 131-147.
- Hartog, J. and Oosterbeek, H.,(1993). "Public and Private Sector Wages in the Netherlands", *European Economic Review*, Vol. 37, pp. 97-114.
- Hou, F., (2004). "The Study on the Return to Investment in Human Capital in Rural China", *Economics Research*, Vol. 12, pp. 75-84.

- Johnson, E. N. and Chow, G. C., (1997). "Rates of Return to Schooling in China", *Pacific Economic Review*, Vol. 2(2), pp. 101-113.
- Lai, D., (1998). "Education, Labor Market and Income Distributions", *Economic Research*, Vol. 5, pp. 42-49.
- Lee, L., (1983). "Generalized Econometric Models with Selectivity", *Econometrica*, Vol. 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*, Vol. 3, pp. 64-76.
- Li, H., (2003). "Economic Transition and Returns to Education in China", *Economics of Education Review*, Vol. 22(3), pp. 317-328.
- Li, S. and Li, W., (1994). "The Estimation of Individual Returns to Education in China's Education Investment", in R. Zhao, S. Li and C. Riskin (eds), *Studies on Income Distribution of Chinese Residents*, China Society Science Press, pp. 442-456.
- Li, S. and Ding, S., (2003). "The Long Term Change in Private Returns to Education in Urban China", *China Society Science*, Vol. 6, pp. 58-72.
- Meng, X., (1998). "Male-female Wage Determination and Gender Wage Discrimination in China's Rural Industrial Sector", *Labour Economics*, Vol. 5(1), pp. 67-89.
- Miller, P.W., (1987). "Gender Differences in Offered and Observed Wages in Canada, 1980", *Canadian Journal of Economics*, Vol. 20, No. 2, pp.225-244.
- Miller, P.W., (2009). "The Gender Pay Gap in the US: Does Sector Make a Difference?", *Journal of Labor Research*, Vol. 30, No. 1, pp.52-74.
- Puhani, P.A., (2000). "The Heckman Correction for Sample Selection and its Critique", *Journal of Economic Surveys*, Vol. 14, No. 1, pp. 53-68.
- 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*.
- Wang, M. and Cai, F., (2008). "Gender Earnings Differential in Urban China", *Review of Development Economics*, Vol. 12(2), pp. 442-454.
- 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*, Vol. 6, pp. 1-7.

- Yu, Xue., (2000). "The Study in Return to Human Capital in Urban Labor Markets during the Transition Period in China", *The Issues of Labor Market and Employment in China*, Southwestern University of Finance and Economics Press, pp.180-195.
- Zhang, J., Zhao Y., Park A. and Song, X., (2005). "Economic Returns to Schooling in Urban China, 1988 to 2001", *Journal of Comparative Economics*, Vol. 33, pp. 730-752.
- Zhao, Y., (1997). "Labor Migration and Returns to Rural Education in China", *American Journal of Agricultural Economics*, Vol. 79(4), pp. 1278-1287.

**APPENDIX**  
**Table 1**  
**Definition of Variables**

Variable	Definition
<b>Dependent Variable</b>	
Log Earnings	Natural logarithm of monthly income.
<b>Independent Variables</b>	
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 <sup>(1)</sup>	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 <sup>(2)</sup>	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 <sup>(3)</sup>	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:* <sup>(1)</sup> West is the benchmark region; <sup>(2)</sup> Collective enterprise is the benchmark ownership; <sup>(3)</sup> The benchmark occupation is unskilled workers (ordinary labourers, service workers, farm workers).

**Table 2**  
**Implied Returns to Education, Adjusting for Effects of ORU Variables,  
Comparisons of Males and Females in Rural Areas**

	% Payoff
Males	3.39
Females	7.68
-no adjustment	
(i) Assuming the same levels of required education, under education and over education as for males	7.13
(ii) as for (i) but also assuming the same earnings effects to required education, under education and over education categories 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.

**Table 3**  
**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 over education as for males	6.86
(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.