The returns to education in rural China: Some new estimates*

Chengfang Liu ⓘ, Ye Li, Shaoping Li, Renfu Luo,
Linxiu Zhang, Scott Rozelle, Spencer Hagist and Jack Hou[†]

We estimate the rates of return to education in rural China using primary survey data collected in 2016. Estimated average returns to education are 3.1 per cent. However, careful statistical analysis is required when estimating the returns to education. The paper demonstrates that when employment interruptions are accounted for, the measured returns to education rise. Our results also confirm that mismeasurement of the wage rate by using an hourly wage rate (versus daily or monthly earnings) raises the estimation of rates of return to education. Finally, our results suggest that the return to education is nonlinear in education levels but only when it reaches the tertiary level.

Key words: experience, mismeasurement, off-farm employment, return to education, rural China.

1. Introduction

Recent research has clearly shown the importance of rural education for growth and development in China. Statistics show that 40 per cent of China's population still permanently resides in rural areas (NBSC 2019), and an even larger percentage (51 per cent) of children aged 0–17 are educated in rural areas (NBSC 2016). As such, the Government has made great efforts to increase investment into rural education in recent decades. For example, China launched the Teacher Incentive Payment Program in 2009 (Loyalka *et al.* 2019), wherein salaries were raised to the levels of civil servants. The central Government invests billions of US dollars each year in the National Teachers Training Program (Li *et al.* 2019). These and other investments and programs are focused on improving education in rural areas.

Despite these efforts, the level of education acquired by China's rural labourers remains low. According to the 2015 micro Census data, only 14 per cent of the labour force residing in rural areas had acquired a senior high school level of education (or higher), compared with 42 per cent of those living in cities and towns (NBSC 2016). Compared with its neighbours, China's labour force has much lower educational attainment. Japan and Korea, for example, have achieved universal senior high school attainment (Chen *et al.* 2017). The proportion of people aged 25–34 who have acquired at least senior high school education is 35 per cent in China, even lower than Indonesia at 45 per cent (OECD 2017).

Why does educational attainment remain at such low levels in rural China? One of the key reasons might be that investments in rural education are perceived to generate relatively low returns. Returns to education are measured as the per cent increase to income gained for each additional year of schooling. Previous studies of China's rural labour force have measured the returns to rural education as consistently low, seldom exceeding 5 per cent (Parish *et al.* 1995; Meng 1996; Johnson and Chow 1997; Ho *et al.* 2002). To the best of our knowledge, only two studies find higher rates of returns. De Brauw and Rozelle (2008) estimated an average return of 6.4 per cent, and Chen *et al.* (2017) estimated a 7.6 per cent average return. In contrast, the global benchmark of returns to education is 9.7 per cent, the average for Asian countries is 9.9 per cent (Psacharopoulos and Patrinos 2004; Patrinos and Psacharopoulos 2010), and the average return in represented developing countries is 7.6 per cent (Peet *et al.* 2015). If such studies are influencing perceptions, and the general public believe that schooling returns in rural areas are low, rural families may be discouraged from keeping their children in school. This assumption could also be used to argue that the Government should focus on other priorities when deciding on public goods expenditures.

While some researchers are puzzled by these low measured returns (De Brauw and Rozelle 2008; Chen *et al.* 2017), a close review of existing literature reveals two limitations that might lead to a consistent underestimation of the returns to education. First and foremost, mismeasurement may be a problem. In fact, there are two different possible forms of mismeasurement. The first may be the choice of what 'wage rate' should be used in studies of returns to education (De Brauw and Rozelle 2008). Previous studies, due to data constraints, mostly rely on daily, monthly, or annual earnings, instead of hourly wage rates, in estimating returns to education (Fleisher and Wang 2005; Meng and Gregory 2005; Li *et al.* 2012). It has been observed that poor people who are less educated are willing to work more hours (Li 2003). As such, when using a monthly wage instead of an hourly wage rate, empirical studies may underestimate returns to education (Li 2003; De Brauw and Rozelle 2008; Chen *et al.* 2017).

The second type of mismeasurement concerns how to account for working experience, especially in cases where experiences are interrupted by time spent on finding new jobs, childbearing or childrearing (Oaxaca 1973; Mincer 1974;

Mincer and Polachek 1980; Albrecht *et al.* 1999; Li 2003; Goldin 2006). Studies have shown that working experience is an important factor affecting the estimate of returns to education (Mincer 1974; Heckman and Li 2003). Since both experience and education affect the measurement of returns and they may be correlated, overestimating experience can lead to an underestimation of the returns to education (Albrecht *et al.* 1999; Li 2003). As such, many studies have taken employment interruptions into considerations, for example, Rimmer and Rimmer (1997), Albrecht *et al.* (1999), Theunissen *et al.* (2011), and Mortelmans and Frans (2017).

Previous studies have identified two channels from which employment interruptions can affect human capital accumulated on the job, which is an important factor in the wage equation (Mincer 1974). In one way, employment interruption will decrease the duration of actual experience, which may influence the accumulation of human capital (Albrecht *et al.* 1999). In another way, the human capital acquired in the past is more likely to be obsoleted when experiencing career break. As such, workers with employment interruptions in the past will find it difficult to re-enter the job market or keeping the previous state (Theunissen *et al.* 2011; Rønsen and Kitterød 2015). Excluding employment interruptions faced by labourers may lead to consistent overestimation of experience and thus underestimate the returns to education.

Due to data constraints, actual work experience is often approximated by potential work experience, as defined by age minus years of schooling and school entry age. Measuring potential work experience this way fails to account for employment interruptions, which makes an inadequate proxy for job-related skills (Mavriplis *et al.* 2010). To the best of our knowledge, few studies on the returns to education in the context of rural China have taken into account employment interruptions when measuring work experience.

The second limitation is treating the returns to education as the same across different education levels (Jamison and Gaag 1987; Byron and Manaloto 1990; Yang 2005; Li *et al.* 2007). As far as we know, the few exceptions include Knight and Song (1991) and Liu (1998), both find that the average annual returns to education at different levels of primary, secondary and tertiary schools are different. Both of these studies, however, were conducted more than two decades ago, long before the expansion of China's education system. By not sufficiently distinguishing different returns to different levels of education, previous studies provide limited evidence for designing relevant education policies.

To fill the gap, this paper seeks to re-estimate returns to education in off-farm employment in rural China. To do so, we must first fix the mismeasurement in working experience caused by ignoring employment interruptions and the mismeasurement in the wage rate caused by ignoring working hours. We further explore the potential nonlinearity in returns to education. With more accurate measures of returns to education, we may

provide empirical evidence for designing relevant education policies in the Chinese context.

This study contributes to the literature in two ways. First, in making our estimate of the rates of return, we correct for both mismeasurements in the wage rate and experience simultaneously. As far as we know, this is the first study that has taken this into account in the context of rural China. Second, we extend the analyses of recent authors by examining the potential nonlinearities in returns to education at different levels (De Brauw and Rozelle 2008; Chen *et al.* 2017).

The rest of the paper is organised as follows. Section 2 describes our data and variables. Section 3 introduces our empirical strategy. Section 4 presents results. The final section concludes the paper.

2. Data

We collected data used in this paper under the China Rural Development Survey (CRDS) project. Starting in 2005, the CRDS is a nationally representative rural survey, containing yearly employment information on individuals in 2,020 rural households. So far, four waves of surveys have been conducted: 2005; 2008; 2012; and 2016. Using a multistage stratified cluster random sampling procedure, the survey covers 25 counties, 50 townships and 101 villages in 5 provinces: Jiangsu; Shaanxi; Sichuan; Jilin; and Hebei. From these villages, a nearly nationally representative sample of 2,020 households was selected.

The CRDS household questionnaire collected detailed information on each household member. If a household member or one of their children was not present, the respondent, who was almost always the household head or spouse, answered. We use the 2016 wave because, among other things, it has the longest employment history information that could be used to measure employment interruptions, a key variable employed in this paper.

In the 2016 wave, in addition to standard demographic information, such as gender and age of each household member, the household questionnaires collected a wide range of employment information. Specifically, trained enumerators asked all household members who were at least 16 years old about their 18-year employment history, from 1998 to 2015, by recall. Extensive pretesting found that the data are fairly accurate.

For each individual in each year, we asked about her/his employment participation. This enables us to identify an individual's working status in each year and then measure the periods of interruption in each respondent employment history. For the purpose of the study, we define an individual as being employment interrupted if he or she has left from job market for at least 1 year between 1998 and 2015. Duration of employment interruption is defined as the accumulative number of years that an individual has left from the job market between 1998 and 2015. If an individual was employed (or working) in any given year, we would ask the specific type of work

performed. To identify whether an individual was working off-farm and whether she/he was a migrant or local employee, we asked her/his place of Household Hukou registration and job location. We also asked whether an individual was self-employed or whether the individual worked for a wage/salary.

We limit the analysis to wage and/or salary workers. This restriction does not exclude agricultural labourers or informal sector employees. However, it does exclude farm and other business owners and the self-employed who do not report wages or salaries but instead report profits or revenues from their operations (which cannot be directly attributed to individuals or to the labour of individuals.

For the purpose of the study, we define an individual's earnings as all income from her/his primary off-farm occupation in 2015. This includes regular monthly wages, bonuses, subsidies and in-kind wages. Moreover, in separate questions about the individual's primary off-farm employment, we asked the number of working hours on an average day, the average number of working days per month and the number of working months in 2015. In this way, we could convert all earnings information into an hourly wage rate. If we do not take working hours into consideration when measuring the wage rate, we may run into underestimation of returns to education in rural China (Card 2001).

The survey also asked detailed information on each person's education. For each household member who ever attended school, we asked the precise age when they started their primary education. We also asked whether they had ever repeated or skipped any grades, or had any sick leaves. For those who finished their education, we further asked the precise age when they left school and finished their education. With this information, we are able to construct two sets of education measures. One is a continuous variable indicating years of completed schooling. The other is a set of dummy variables indicating whether the highest grade an individual successfully completed is at the primary level (1–6 years), junior high level (7–9 years), senior high level (10–12 years) or tertiary level (13 years and above).

With detailed information on each person's education and her/his participation in employment between 1998 and 2015, we are able to construct two proxies of potential experience. The first is *unadjusted potential experience*. We follow the standard practice in the literature and define it as the precise age in 2015 minus the number of years of full-time schooling, minus the precise age when the individual started primary education. The other is *adjusted potential experience*. We take employment interruptions into account and define it as unadjusted potential experience minus the duration of employment interruptions between 1998 and 2015.[1] In other words:

---

[1] Our data show the correlation between unadjusted potential experience and age is 0.98, whereas it is 0.96 between adjusted potential experiences and age. The correlation between unadjusted and adjusted potential experiences is 0.98.

$$\text{Unadjusted potential experience} = \text{Age} - \text{Years of schooling} \\ - \text{Precise school entry age.} \tag{1}$$

$$\text{Adjusted potential experience} = \text{Age} - \text{Years of schooling} \\ - \text{Precise school entry age} \tag{2} \\ - \text{Years of employment interruptions.}$$

For the purpose of the study, we include sample individuals based on the following criteria. First, we include those who were 16–65 years old in 2015. Since rural labourers often keep working after they reach the mandatory retirement age, we chose 65 as the upper bound. The minimum legal age for employment, 16, was chosen as the lower bound (De Brauw and Rozelle 2008). Second, we exclude those who had children born before 1998 since we do not have information on their employment interruptions. Thus, we ended up with 5,681 individuals, the sample for this paper. Table 1 depicts the sample distribution.

3. Empirical model

Following prior literature (Mincer 1974; Willis 1986), we model the effect of education on earnings with a standard Mincerian wage equation as follows:

$$\ln Y_i = \alpha_1 + \beta S_i + \alpha_2 E_i + \alpha_3 E_i^2 + \gamma Z' + \varepsilon_i, \tag{3}$$

where Y_i denotes the hourly wage rate for individual i. S refers to years of schooling. E refers to potential experience and E^2 its quadratic term. α_1 is the constant term which denotes the base hourly wage rate without any full-time schooling. The coefficient of interest, β, describes the percentage change in earnings due to 1-year marginal change in attained schooling, S. Z' represents a vector of covariates that includes the following six sets: (i) a dummy variable to indicate whether the individual is a male or female (De Brauw and

Table 1 Sample distribution

	Number of observations	Percentage (%)
Whole sample	5,681	100
of which:		
Jiangsu	1,099	19.35
Sichuan	1,124	19.79
Shaanxi	1,315	23.15
Jilin	1,083	19.06
Hebei	1,060	18.66

Source: Authors' survey.

Rozelle 2008); (ii) a dummy variable to indicate whether an individual has skills which may account for wage premiums (Chen *et al.,* 2017)[2]; (iii) a dummy variable to specify whether the individual is married or not, since marital status has been found to influence individuals' decisions vis-à-vis different types of jobs and wage rates (Huffman and Lange 1989); (iv) a dummy variable to control for the possibility that migrants earn more than local wage earners due to local supply/demand conditions (Maurer-Fazio and Dinh 2004; De Brauw and Rozelle 2008); (v) four occupational dummy variables to indicate different occupational types with manual worker as the base[3]; and (vi) four provincial dummy variables to control for different base wage rates in different provinces with Jiangsu province as the base. The standard errors are clustered at the village level.

3.1 Accounting for potential endogeneity

There has been a concern that the ordinary least-squares (OLS) estimate of returns to education could be biased due to endogeneity from at least two sources. One source of endogeneity is self-selection. De Brauw and Rozelle (2008) point out that rural labourers may decide whether they enter the off-farm labour market in a non-random manner. The other source of endogeneity is omitted variables. Some scholars caution that individual traits such as ability related to years of schooling as well as income may confound the results from the Mincerian wage equation (Griliches 1977; Willis 1986).

Being aware of the potential endogeneity, we apply the standard Heckman two-step model to yield a more consistent estimation (Ashenfelter and Heckman 1974). Following the literature, we specify six variables as the exclusions: the number of male household members aged 16–65; the number of female household members aged 16–65; per capita landholding in mu; the value of household assets in the log value of durables in yuan in 2015; the number of children under 6 years old; and the number of elderly people over 70 years old (De Brauw and Rozelle 2008; Zhou 2012).[4]

3.2 Accounting for nonlinearity and heterogeneity in the returns to education

In addition to endogeneity, we need to address another concern. The Mincer equations have been criticised for ignoring the potential nonlinearity in returns to education by treating the returns as the same for each additional year of education (Card 2001; Li 2003). We take two steps to address this concern. First, we substitute the years of schooling variable with dummy

---

[2] An individual is defined as 'having skills' if she or he has acquired any ability that allows her or him to do things that would bring her or him any money or in-kind benefits, such as driving, making clothes, building houses, repairing cars, etc.

[3] We organise the 38 occupations in the household questionnaire into five major categories: professional; leader of entities; clerical staff; service worker; and manual worker.

[4] 1 mu is equivalent to 0.165 acres; 1 yuan is equivalent to 0.145 US Dollars.

variables to indicate an individual's level of education. Second, we separate years of schooling into two groups, compulsory schooling and postcompulsory schooling.

There is still another concern we need to address, which is the potential heterogeneity in the returns to education by occupation types. The effect of employment interruption on wage rage may vary across occupation types (Adda *et al.* 2017). For example, for service and manual work, a few years of employment interruptions may play a minimal role in affecting their hourly wage rate, while the effects can be significant for professionals and clerical staffs. To address this concern, we include the interaction terms of the occupation dummies with the duration of employment interruptions into the Mincerian wage equation.

4. Result

4.1 Descriptive result

Descriptive statistics generated from our data show that the profile of sample individuals is fairly typical of people from rural areas. The sex ratio of our sample is 104, which is exactly the same as reported for rural areas by the Sixth National Population Census (NBSC 2011). Similarly, 33 per cent of our sample individuals are aged 35 and under and 79 per cent have attained junior high school or below education, which are almost the same as those reported by the Census (Table 3).

We now turn to a description of our key variables of interest. Among the 5,681 observations, 21 per cent have had some form of employment interruptions. On average, the employment interruption lasts 5.92 years (Table 2). For those individuals with employment interruptions, the adjusted potential experience (when considering interruption) is 21.97 years on average, much less than the unadjusted potential experience (when not considering interruption) at 27.90 years (Table 3, Rows 7–8, Column 3). The average years of schooling is 7.66, less than the 9-year compulsory education (Row 2, Column 1). The average hourly wage rate is 15.28 yuan per hour (Row 1, Column 1).

Table 3 also compares labourers with employment interruptions to those without employment interruptions. Our data show that those with interruptions tend to be female and younger, have lower off-farm wages, are less

Table 2   Summary statistics of employment interruption

	Mean
% with interruptions	20.68
of those with interruptions: Duration of interruptions (years)	5.92

Source: Authors' survey.

Table 3 Summary statistics of variables

	Variables	All sample		With interruption		Without interruption		Difference test (*P*-value)
		Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)	H_0: (3) = (5) (7)
(1)	Off-farm wage rate (Yuan/hour)	15.28	15.73	12.59	13.27	15.64	15.99	0.002**
Education:								
(2)	Years of schooling	7.66	3.94	7.37	3.86	7.74	3.96	0.004**
(3)	Primary and below (1 = yes)	0.26	0.44	0.25	0.43	0.26	0.44	0.366
(4)	Junior high (1 = yes)	0.43	0.49	0.45	0.50	0.42	0.49	0.072*
(5)	Senior high (1 = yes)	0.15	0.36	0.15	0.36	0.16	0.36	0.575
(6)	Tertiary (1 = yes)	0.06	0.24	0.03	0.18	0.07	0.25	0.000***
Potential experience								
(7)	Unadjusted (year)	29.76	15.78	27.90	16.92	30.24	15.44	0.000***
(8)	Adjusted (year)	28.53	15.96	21.97	16.26	30.24	15.44	0.000***
Occupation								
(9)	Cadre (1 = yes)	0.16	0.37	0.12	0.33	0.17	0.21	0.000***
(10)	Professional (1 = yes)	0.01	0.10	0.01	0.10	0.01	0.09	0.933
(11)	Clerical staff (1 = yes)	0.04	0.19	0.01	0.10	0.17	0.38	0.000***
(12)	Service worker (1 = yes)	0.08	0.27	0.04	0.19	0.09	0.29	0.000***
(13)	Manual worker (1 = yes)	0.27	0.44	0.31	0.34	0.13	0.46	0.000***
Other individual characteristics								
(14)	Age (year)	43.42	13.50	41.27	14.75	43.98	13.10	0.000***
(15)	Gender (1 = male)	0.51	0.49	0.27	0.44	0.57	0.49	0.000***
(16)	Skill training (1 = yes)	0.35	0.48	0.23	0.42	0.39	0.49	0.000***
(17)	Marital status (1 = married)	0.89	0.31	0.90	0.31	0.89	0.331	0.705
(18)	Migrant (1 = yes)	0.36	0.48	0.22	0.42	0.39	0.49	0.000***
Household characteristics								
(19)	Labour endowment: male (persons)	1.89	0.71	1.90	0.74	1.89	0.71	0.602
(20)	Labour endowment: female (persons)	1.84	0.73	1.95	0.74	1.81	0.73	0.000***
(21)	Household landholding (mu)	7.42	10.54	5.60	7.44	7.89	11.16	0.000***
(22)	Assets holding (10,000 yuan) in log	2.39	1.68	2.44	1.86	2.37	1.63	0.242
(23)	Number of children under 6 years old (persons)	0.80	0.97	0.90	0.98	0.77	0.97	0.000***
(24)	Number of elderly people over 70 years old (persons)	0.54	0.82	0.48	0.78	0.56	0.82	0.003**
	Observations	5,681		1,175		4,506		

Note: ***Significant at 1% level; **significant at 5% level; *significant at 10% level.
Source: Authors' survey.

educated and less experienced, are more likely to work as a cadre and less likely as a clerical staff or manual worker and are less likely to have migrated or have had any skill training. Those with more interruptions also have more children under 6 years old to take care of and fewer elderly people over 70 years old in the home. These results provide suggestive evidence that more educated and skilled labourers are less likely to get interrupted in their off-farm employment, which lends more justification for the necessity to take self-selection bias into consideration when estimating the returns to education.

4.2 Result from multivariate analyses

Table 4 reports the findings from the regressions of log hourly wage rate on education, potential experience and control variables. For the sake of clarity, we only report the coefficient on the education variable.[5] Panel A shows results from OLS regressions, whereas Panel B shows results from the Heckman two-step regressions accounting for self-selection bias. Column 1 shows results from regressions when using years of schooling to measure education, thereby neglecting the potential nonlinearity in the returns to education. Columns 2–4 report results from regressions when using three dummy variables to indicate the level of highest grade completed with primary level education as the base. Rows 1 and 4 show results from regressions when using unadjusted potential experience. Rows 2 and 5 report results from regressions when using adjusted potential experience. Rows 3 and 6 report percentage of underestimation of returns to education caused by using unadjusted experience compared to using adjusted experience, holding other things the same.

Results from the multivariate regression estimates consistently show that mismeasurement in potential experiences due to ignoring employment interruptions does lead to underestimation of returns to education by 0.2–0.9 percentage points, depending on the estimation methods and education measures. Specifically, with years of schooling as a measure of education, the OLS model using unadjusted potential experience yields a 2.1 per cent return to a year of schooling (Table 4, Row 1, Column 1). When using the adjusted potential experience, the return to education is slightly higher by 0.3 percentage point (at 2.4 per cent) (Row 2, Column 1). These estimations are lower than those obtained by using the hourly wage rate in rural China in previous studies (Li 2003; De Brauw and Rozelle 2008).

Our data also show that return to education is nonlinear but not only when it reaches the tertiary level (Rows 1–2, Columns 2–4). When education-level dummy variables are used, using primary education and below as the base, we find the nonlinearity in returns to education does not take effect until it

---

[5] Regression results from Heckman two step are shown in Table A1 in Appendix. Full regression results of each specification are available upon request to the corresponding author.

Table 4 Estimation of return of education

Proxy for potential experience	Dependent variable: log of hourly wage rate			
	Average returns to education (1)	Returns to junior high education (2)	Returns to Senior high education (3)	Returns to tertiary education (4)
OLS				
(1) Unadjusted	0.021***	−0.015	0.022	0.352***
(2) Adjusted	0.024***	0.022	0.024	0.356***
(3) Difference (= 1–2)	−0.003	/	/	−0.004
Heckman two step				
(4) Unadjusted	0.029***	−0.001	0.047	0.394***
(5) Adjusted	0.031***	0.008	0.056	0.403***
(6) Difference (= 4–5)	−0.002	/	/	−0.009

Note: ***Significant at 1% level; **significant at 5% level; *significant at 10% level. /, not applicable; OLS, ordinary least-squares.
Source: Authors' survey.

reaches the level of tertiary education. When not adjusting potential experience, those with junior and senior high education earn statistically the same as those with primary education and below (Row 1, Columns 2–3). However, those with tertiary education earn 35.2 per cent more (Row 1, Column 4). When using adjusted potential experience, the same pattern holds and the estimated returns to tertiary education are higher by 0.4 percentage point (at 35.6 per cent). These numbers are higher than those reported by previous studies. Liu (1998), for example, reports that wages for university graduates are only 30 per cent higher than for those who finished primary school. Knight and Song (1991) find that a university or college graduate receives, on average, only 10 per cent more than someone with a primary school education or less. This might be explained by the overall improvement of rural education and by our corrected measurement methods in potential experiences, or by the fact that the real returns to college have risen as automation and globalisation have lowered return to low wage, unskilled employment.

As a robustness check, we also estimate results using the Heckman two-step model.[6] In general, the results from Heckman two-step model are consistent with those from the OLS but its magnitudes of the returns to education are bigger. The return to education, with years of schooling as a measure of education, is 2.9 per cent with unadjusted potential experience, whereas it is 3.1 per cent with adjusted potential experience, an underestimate by 0.2 percentage point (Table 4, Rows 4–6, Column 1). When we replace the years of schooling for the education-level dummy variables, our data show that when using unadjusted potential experience, the returns to the junior

---

[6] We also found a higher rate of return of education among younger workers than their older counterparts, which is similar to de Brauw and Rozelle (2008).

high or senior high level are not significantly different, whereas the returns to tertiary educations are 39.4 per cent higher than the primary education. When using adjusted potential experience, the same pattern holds, and the estimated returns to tertiary education are higher by 0.9 percentage point (at 40.3 per cent) (Rows 4–6, Columns 2–4).

Our results from Heckman two-step model provide evidence of heterogeneity in returns to education by types of occupation. When we include the interaction terms of occupation dummies with the duration of employment interruptions into the equation and re-estimate it, our results show that the effect of employment interruption on wage rate is different between service workers and other employment types. Nonetheless, the estimated coefficient on the returns to education remains substantially the same (3.1 per cent) as when occupational dummies nor their interactions were included (Table A2).

In addition to heterogeneity in returns to education by types of occupation, our data also provide evidence that the effect of employment interruption on wage rate varies by province. Specifically, when we measure education by years of schooling, the estimated coefficients on education are significant in four out of the five by province regressions. The return to education is the highest in Hebei province, followed by Jilin, Shaanxi and Jiangsu. When we measure education by dummies for education levels, the estimated coefficient on junior high education dummy is not significant in any of the five regressions whereas it is significant in only one of the five regressions when we measure education by the senior high education dummy. In contrast, when it comes to tertiary education level, the estimated coefficients are statistically significant in all of the five regressions, with Jilin being the highest, followed in turn by Sichuan, Hebei, Shaanxi and Jiangsu (Table A3).

4.3 Reconciling with other studies in rural China

As mentioned above, some previous research may have generated biased estimates of returns to education due to mismeasurements in potential experiences and wage rate. Most previous studies use earnings as their dependent variable without considering hours worked, due to unavailability of data (Gregory and Meng 1995; Parish *et al.* 1995; Meng 1996; Ho *et al.* 2002; Chen *et al.* 2017). To the best of our knowledge, no study yet has taken employment interruptions and hours worked simultaneously into account when estimating returns to education (see Table A4). Therefore, in order to reconcile our estimates for the returns to education among labourers in rural China with other authors' findings, we follow wage and experience measurements from the existing literature to re-estimate our results.

We use different measurements of wage and potential experience in the reconciling exercise based on the OLS specification (Table 5). We find that mismeasurement of wage rate and potential experience accounts for a large portion of the gap between our estimations and those from previous studies. Specifically, in our study, we replace wage measurement with daily wage rate,

monthly wage rate and annual earning, separately. This allows us to examine the effect of not correcting the mismeasurements of wage and potential experience. Compared with our estimate of 3.1 per cent based on hourly wage rate and adjusted potential experience, we found that using the daily wage rate led the returns to education to be underestimated, by 0.2 percentage points, to 2.9 per cent, regardless of the measurement of potential experience. Moreover, we also find that using annual earnings led the returns to education to be underestimated by 1.0 percentage point (using unadjusted potential experience) or 0.7 percentage point (using adjusted potential experience). These underestimations are all statistically significant at least at 10 per cent level. This further lends evidence, as we predicted, to the idea that mismeasuring the wage rate and potential experience tends to underestimate returns to education. In contrast, when using the hourly wage rate with unadjusted potential experience, although the point estimate of returns to education tends to be larger than when using adjusted potential experience, the difference is not statistically significant. Similarly, when using monthly earnings with either unadjusted or adjusted potential experiences, the point estimates are larger but not statistically significant.

5. Conclusion

In this paper, we have used primary survey data of rural labour force to re-estimate the returns to education in rural China. Correcting for self-selection bias in off-farm employment and for the measurement of wage rate and

Table 5 Reconciling returns to education in rural China

Measurement of wage rate	Return to education	Difference compared to 3.1% (aka the returns to education using hourly wage rate and adjusted experience)	Difference test H_0: (2) = 3.1% (*P*-value)
	(1)	(2)	(3)
(1) Hourly wage rate, unadjusted experience	2.9%	0.2%	0.37
(2) Daily earning, unadjusted experience	1.7%	1.4%	0.03*
(3) Daily earning, adjusted experience	1.7%	1.4%	0.04*
(4) Monthly earning, unadjusted experience	2.1%	1.0%	0.87
(5) Monthly earning, adjusted experience	2.1%	1.0%	0.88
(6) Annual earning, unadjusted experience	2.1%	1.0%	0.02*
(7) Annual earning, adjusted experience	2.4%	0.7%	0.01**

**Significant at 5% level; *significant at 10% level.
Source: Authors' survey.

potential experience, we find that among rural labourers engaged in off-farm work, the mean return to education is 2.4 per cent (OLS) or 3.9 per cent (Heckman two-step) in terms of their hourly wage rate. This estimate is lower than many previously reported estimates of the returns to education in China (Parish *et al.* 1995; Johnson and Chow 1997; Li 2003; De Brauw and Rozelle 2008; Chen *et al.* 2017), but higher than others (Meng 1996; Yang 1997; Ho *et al.* 2002).

Our study also shows the importance of using the hourly wage rate rather than daily, monthly or annual earnings. Because highly educated workers and workers with more wealth or other endowments tend to assign different values to work and leisure, they may choose different amounts of time to work (Card 2001; De Brauw and Rozelle 2008; Chen *et al.* 2017). Using the hourly wage rate can therefore control for the time a person chooses to work. After replacing the hourly wage rate using the annual wage rate, monthly wage rate and daily wage rate, we find that the returns to education see a decrease of 0.7–1.4 percentage points. Therefore, not using the hourly wage rate can generate a non-negligible underestimation of returns to education. At the same time, it is important to pay attention to the potential self-selection bias. The Heckman two-step model suggests that the result is consistent with our previous findings. Finally, we find that returns to education are nonlinear only until they reach the tertiary level. A tertiary level of education means larger wage premiums.

Even after addressing the mismeasurements of wage rate and experience as well as the self-selection of earlier studies, this paper finds that return to rural schooling in China still remains low. This may explain why many rural households decide to allow their children to leave school so early (Mo *et al.* 2013; Yi *et al.* 2015). In fact, in one study, it was estimated that the cumulative dropout rate across all windows of secondary education may be as high as 63 per cent (Shi *et al.* 2015). One possible reason why the return to rural schooling in China is so low is that the rural education quality is bad. Indeed, lots of students reported that they regret to go to vocational high school (Shi *et al.* 2015) because those school qualities are too bad, and the stories of these students are verified in Loyalka *et al.* (2015). Therefore, improving rural education quality to let rural children obtain the skills they need to participate in China's rapidly changing, dynamic, increasingly high-tech/high skill economy is very important.

Finally, our estimates also suggest that increasing tertiary education availability may help rural households. We find that the rates of returns to tertiary education are approximately 40 per cent which is higher than those found in previous studies. Given these high returns, the Chinese Government should expand its higher education system. As the Chinese economy is undergoing transition from high-speed growth to high-quality development, it is essential to improve the quality of its human capital.

Our study is not without limitations. Even with the Heckman two-step, we are not able to fully address the potential endogeneity, especially the

endogeneity due to omitted ability variable bias. We tried to follow the literature (Ashenfelter and Zimmerman 1997; Brown 2006) and use father's education as a proxy for the individual's ability. Unfortunately, as the fathers of many sample individuals do not belong to our sample subjects, we obtained father's education information for a subsample of 1,171 individuals. Regression results from this subsample are consistent with the results that we have shown above, mismeasurement in potential experience and wage rate would underestimate the returns to education, and the return to education is nonlinear in education levels but only when it reaches the tertiary level.

Nevertheless, measures should be taken to help rural labourers, especially those who have employment interruptions, perform well in the workforce. To meet this goal, the Government needs to prioritise investment in rural education, especially for those with employment interruptions. Given limited access to public services of childcare and elderly care, people in rural areas are often tasked with childrearing and taking care of elderly people, so they are more likely to suffer from job interruptions. When people leave the job market, their human capital will depreciate. The longer the interruption, the more difficult their return to the labour market. Therefore, investment in rural centres for childcare and nursing of the elderly is needed to help relieve this burden on the rural labour force. Moreover, governments should provide professional retraining programs to those who have left the job market for a long time. Taking such measures, which are less costly than more typical welfare measures (which must be carried out indefinitely) will serve the dual purpose of expediting growth in rural areas while also correcting for the undue stress currently saddled on rural labourers.

References

Adda, J., Dustmann, C. and Stevens, K. (2017). The career costs of children, *Journal of Political Economy* 125, 293–337.

Albrecht, J.W., Edin, P.A., Sundström, M. and Vroman, S.B. (1999). Career interruptions and subsequent earnings: a reexamination using Swedish data, *Journal of Human Resources* 34, 294–311.

Ashenfelter, O. and Heckman, J. (1974). The estimation of income and substitution effects in a model of family labor supply, *Econometrica* 42, 73–85.

Ashenfelter, O. and Zimmerman, D.J. (1997). Estimates of the returns to schooling from sibling data: fathers, sons, and brothers, *Review of Economics and Statistics* 79, 1–9.

Brown, P.H. (2006). Parental education and investment in children's human capital in rural China, *Economic Development and Cultural Change* 54, 759–789.

Byron, R.P. and Manaloto, E.Q. (1990). Returns to education in China, *Economic Development and Cultural Change* 38, 783–796.

Card, D. (2001). Estimating the return to schooling: progress on some persistent econometric problems, *Econometrica* 69, 1,127–1,160.

Chen, Q., Xu, J., Zhao, J. and Zhang, B. (2017). Endogenous schooling, school proximity and returns to rural schooling in northwestern China, *China Agricultural Economic Review* 9, 270–286.

De Brauw, A. and Rozelle, S. (2008). Reconciling the returns to education in off-farm wage employment in rural China, *Review of Development Economics* 12, 57–71.

Fleisher, B.M. and Wang, X. (2005). Returns to schooling in China under planning and reform, *Journal of Comparative Economics* 33, 265–277.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family, *American Economic Review* 96, 1–21.

Gregory, R.G. and Meng, X. (1995). Wage determination and occupational attainment in the rural industrial sector of China, *Journal of Comparative Economics* 21, 353–374.

Griliches, Z. (1977). Estimating the returns to schooling: some econometric problems, *Econometrica* 45, 1–22.

Heckman, J.J. and Li, X. (2003). Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000, *Pacific Economic Review* 9, 155–171.

Ho, S.P.S., Dong, X., Bowles, P. and MacPhail, F. (2002). Privatization and enterprise wage structures during transition: evidence from rural industry in China, *The Economics of Transition* 10, 659–688.

Huffman, W.E. and Lange, M.D. (1989). Off-farm work decisions of husbands and wives: joint decision making, *The Review of Economics and Statistics* 71, 471–480.

Jamison, D.T. and Gaag, J.V.D. (1987). Education and earnings in the People's Republic of China, *Economics of Education Review* 6, 161–166.

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

Knight, J.B. and Song, L. (1991). The determinants of urban income inequality in China, *Oxford Bulletin of Economics and Statistics* 53, 123–154.

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

Li, H., Liu, P.W., Zhang, J. and Ma, N. (2007). Economic returns to communist party membership: evidence from urban Chinese twins, *The Economic Journal* 117, 1,504–1,520.

Li, H., Liu, P.W. and Zhang, J. (2012). Estimating returns to education using twins in urban China, *Journal of Development Economics* 97, 494–504.

Li, G., Loyalka, P.K., Popova, A. and Shi, Z. (2019). Does teacher training actually work? Evidence from a large-scale randomized evaluation of a National Teacher Training Program, *American Economic Journal: Applied Economics* 11, 128–154.

Liu, Z. (1998). Earnings, education, and economic reforms in urban China, *Economic Development and Cultural Change* 46, 697–725.

Loyalka, P., Huang, X., Zhang, L., Wei, J., Yi, H., Song, Y., Shi, Y. and Chu, J. (2015). The impact of vocational schooling on human capital development in developing countries: evidence from China, *The World Bank Economic Review* 30, 1–28.

Loyalka, P.K., Sylvia, S., Liu, C., Chu, J. and Shi, Y. (2019). Teacher performance pay design and the distribution of student achievement, *Journal of Labor Economics* 37, 621–662.

Maurer-Fazio, M. and Dinh, N. (2004). Differential rewards to, and contributions of, education in urban China's segmented labor markets, *The Pacific Economic Review* 9, 173–189.

Mavriplis, C., Heller, R., Beil, C., Dam, K., Yassinskaya, N., Shaw, M. and Sorensen, C. (2010). Mind the gap: women in STEM career breaks, *Journal of Technology Management & Innovation* 5, 140–151.

Meng, X. (1996). An examination of wage determination in China's rural industrial sector, *Applied Economics* 28, 715–724.

Meng, X. and Gregory, R.G. (2005). Intermarriage and the economic assimilation of immigrants, *Journal of Labor Economics* 23, 135–174.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Columbia University Press, New York, NY.

Mincer, J. and Polachek, S.W. (1980). Family investments in human capital: earnings of women, *The Economics of Women & Work* 82, 76–76.

Mo, D., Zhang, L., Yi, H., Luo, R., Rozelle, S. and Brinton, C. (2013). School dropouts and conditional cash transfers: evidence from a randomised controlled trial in rural China's junior high schools, *Journal of Development Studies* 49, 190–207.

Mortelmans, D. and Frans, D. (2017). Wage differentials after a career break: a latent growth model on Belgian register data, *Longitudinal and Life Course Studies* 8, 169–190.

NBSC (National Bureau of Statistics of China) (2011). *Tabulation on the 2010 Population Census of the People's Republic of China*. China Statistics Press, Beijing.

NBSC (National Bureau of Statistics of China) (2016). *The 2015 1% National Population Sample Survey of the People's Republic of China*. China Statistics Press, Beijing.

NBSC (National Bureau of Statistics of China) (2019). *China Statistical Yearbook 2018*. China Statistics Press, Beijing.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets, *International Economic Review* 14, 693–709.

OECD (2017). *Education at a Glance 2017: OECD Indicators*. OECD Publishing, Paris.

Parish, W.L., Zhe, X. and Li, F. (1995). Nonfarm work and marketization of the Chinese countryside, *The China Quarterly* 143, 697–730.

Patrinos, H.A. and Psacharopoulos, G. (2010). Returns to education in developing countries, in Peterson, P., Baker, E. and McGaw, B. (eds), *International Encyclopedia of Education* (3rd edn). Elsevier, Amsterdam, pp. 305–312.

Peet, E.D., Fink, G. and Fawzi, W. (2015). Returns to education in developing countries: evidence from the living standards and measurement study surveys, *Economics of Education Review* 49, 69–90.

Psacharopoulos, G. and Patrinos, H.A. (2004). Returns to investment education: a further update, *Education Economics* 12, 111–134.

Rimmer, R. and Rimmer, S. (1997). Employment breaks, pay, and career development among Australian women, *Gender, Work & Organization* 4, 202–217.

Rønsen, M. and Kitterød, R.H. (2015). Gender-equalizing family policies and mothers' entry into paid work: recent evidence from Norway, *Feminist Economics* 21, 59–89.

Shi, Y., Zhang, L., Ma, Y., Yi, H., Liu, C., Johnson, N., Chu, J., Loyalka, P. and Rozelle, S. (2015). Dropout in rural China's secondary schools: a mixed-methods analysis, *China Quarterly* 224, 1,048–1,069.

Theunissen, G., Verbruggen, M., Forrier, A. and Sels, L. (2011). Career sidestep, wage setback? The impact of different types of employment interruptions on wages, *Gender, Work & Organization* 18, e110–e131.

Willis, R.J. (1986). Wage determinants: a survey and reinterpretation of human capital earnings functions, in Ashenfelter, O. and Layard, R. (eds), *Handbook of Labor Economics, I*. Elsevier Science Publishers BV, Amsterdam, pp. 525–602.

Yang, D.T. (1997). Education and off-farm work, *Economic Development and Cultural Change* 45, 613–632.

Yang, D.T. (2005). Determinants of schooling returns during transition: evidence from Chinese cities, *Journal of Comparative Economics* 33, 244–264.

Yi, H., Song, Y., Liu, C., Huang, X., Zhang, L., Bai, Y., Ren, B., Shi, Y., Loyalka, P., Chu, J. and Rozelle, S. (2015). Giving kids a head start: the impact and mechanisms of early commitment of financial aid on poor students in rural China, *Journal of Development Economics* 113, 1–15.

Zhou, C. (2012). Gender differences in rural off-farm labour supply in economically developed areas, *Issues in Agricultural Economy* 3, 43–49 (in Chinese).

Appendix

Table A1 Results from Heckman two step: Estimation of returns to education

		(1)	(2)	(3)	(4)
Education					
(1)	Years of schooling	0.029*** (0.006)	0.031*** (0.031)	NA	NA
(2)	Junior high education (1 = yes)	NA	NA	−0.001 (0.041)	0.008 (0.040)
(3)	Senior high education (1 = yes)	NA	NA	0.047 (0.053)	0.056 (0.053)
(4)	Tertiary education (1 = yes)	NA	NA	0.394*** (0.074)	0.403*** (0.073)
Potential experience					
(5)	Unadjusted (year)	0.023*** (0.006)	NA	0.026*** (0.006)	NA
(6)	Unadjusted, squared (year^2)	−0.000*** (0.000)	NA	−0.001*** (0.000)	NA
(7)	Adjusted (year)	NA	0.026*** (0.005)	NA	0.028*** (0.005)
(8)	Adjusted, squared (year^2)	NA	−0.000*** (0.000)	NA	−0.001*** (0.000)
Occupation					
(9)	Cadre (1 = yes)	−0.186*** (0.033)	−0.184*** (0.033)	−0.174*** (0.033)	−0.172*** (0.033)
(10)	Professional (1 = yes)	−0.115 (0.097)	−0.115 (0.098)	−0.135 (0.095)	−0.134 (0.095)
(11)	Clerical staff (1 = yes)	0.126** (0.057)	0.124** (0.057)	0.080 (0.057)	0.079 (0.057)
(12)	Service worker (1 = yes)	−0.212*** (0.053)	−0.215*** (0.053)	−0.218*** (0.054)	−0.220*** (0.053)
Other variables					
(13)	Gender (1 = male)	0.176*** (0.030)	0.169*** (0.031)	0.206*** (0.031)	0.199*** (0.031)
(14)	Skill training (1 = yes)	0.317*** (0.035)	0.323*** (0.036)	0.271*** (0.035)	0.275*** (0.035)
(15)	Marital status (1 = married)	0.043 (0.059)	0.018 (0.058)	0.073 (0.056)	0.059 (0.055)
(16)	Migrant (1 = yes)	0.618*** (0.181)	0.714*** (0.182)	0.380** (0.170)	0.446*** (0.172)
(17)	Province Dummies	YES	YES	YES	YES
(18)	Constant	1.259*** (0.256)	1.133*** (0.253)	1.700*** (0.221)	1.630*** (0.220)
(19)	Inverse Mills Ratio	0.402** (0.171)	0.492*** (0.172)	0.185*** (0.061)	0.246** (0.063)
(20)	Observations	5,681	5,681	5,681	5,681

Note: ***Significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors clustered in village level are in parentheses. NA, not applicable.
Source: Authors' survey.

Table A2 Results from Heckman two step with occupational dummies and interaction terms

		(1)
Education		
(1)	Years of schooling	0.031*** (0.007)
Potential experience		
(2)	Adjusted (year)	0.024*** (0.005)
(3)	Adjusted, squared (year^2)	−0.000*** (0.000)
Occupation		
(4)	Cadre (1 = yes)	0.209*** (0.054)
(5)	Professional (1 = yes)	0.131** (0.058)
(6)	Clerical staff (1 = yes)	−0.167 (0.102)
(7)	Service worker (1 = yes)	0.178*** (0.034)
Interruption and interactions		
(8)	Years of interruption (year)	−0.013 (0.012)
(9)	Years of interruption × Cadre	−0.012 (0.036)
(10)	Years of interruption × Professional	−0.117 (0.126)
(11)	Years of interruption × Clerical staff	−0.003 (0.015)
(12)	Years of interruption × Service worker	−0.091* (0.053)
Other variables		YES
Province dummies		YES
Observations		5,652

Note: ***Significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors clustered in village level are in parentheses. We also control for experience, occupation types, gender, skill training, marital status and migrant in the estimations.
Source: Authors' survey.

Table A3 Results from Heckman two step for subsample by province

	Returns to years of education	Returns to junior high education	Returns to senior high education	Returns to tertiary education
Jiangsu	0.022** (0.011)	−0.032 (0.076)	−0.061 (0.094)	0.241* (0.135)
Sichuan	0.017 (0.014)	−0.020 (0.079)	0.042 (0.111)	0.483*** (0.169)
Shaanxi	0.023* (0.012)	0.073 (0.087)	0.110 (0.115)	0.351** (0.144)
Jilin	0.061** (0.024)	0.138 (0.142)	0.361** (0.184)	0.558** (0.243)
Hebei	0.256*** (0.074)	−0.073 (0.077)	−0.090 (0.099)	0.402*** (0.139)
Observations	5,681	5,681	5,681	5,681

Note: ***Significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors clustered in village level are in parentheses. We also control for experience, occupation types, gender, skill training, marital status and migrant in the estimations.
Source: Authors' survey.

Table A4 Existing literature estimating returns to education in China

Authors (date)	Study area	Study period	Measurement of wage rate	Measurement of experience	Methodology	Average returns to Education
Gregory and Meng (1995)	Rural	1985	Annual earning	Unadjusted experience	OLS	0
Parish *et al.* (1995)	Rural	1993	Annual earning	Unadjusted experience	OLS	3.1%
Meng (1996)	Rural	1986–1987	Daily earning	Unadjusted experience	OLS	0.7–1.1%
Johnson and Chow (1997)	Rural	1988	Monthly earning	Unadjusted experience	OLS	4.02%
Yang (1997)	Rural	1990	Daily earning	Unadjusted experience	OLS	2.3%
Ho *et al.* (2002)	Rural	1998	Annual earning	Unadjusted experience	OLS	3.2–5.0%
De Brauw and Rozelle (2008)	Rural	2002	Hourly wage rate	Unadjusted experience	OLS/ Heckman	3.3–6.0%
Chen *et al.* (2017)	Rural	2004	Daily earning	Unadjusted experience	OLS/2SLS	5.6–7.6%

Note: OLS, ordinary least-squares.
Source: Compiled by the authors.