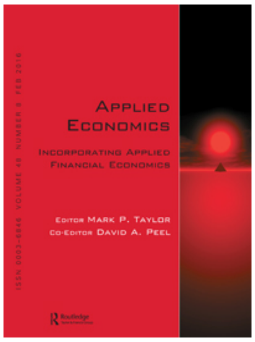


Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

Returns to schooling for urban and migrant workers in China: a detailed investigation

Chris Sakellariou & Zheng Fang

APPLIED ECONOMICS, 2016
VOL. 48, NO. 8, 684–700
http://dx.doi.org/10.1080/00036846.2015.1085642

Routledge
Taylor & Francis Group

Returns to schooling for urban and migrant workers in China: a detailed investigation

Chris Sakellariou and Zheng Fang

Division of Economics, Humanities and Social Sciences, Nanyang Technological University, Singapore, Singapore

ABSTRACT
We use a new data set, the 2009 Rural Urban Migration in China (RUMiC) to estimate returns to schooling in China using instrumental variable (IV) estimation. After identifying a set of instruments, we conduct comprehensive validity and relevance testing of different combinations of instruments as well as robustness analysis of our estimates for rural-to-urban migrants and urban residents in China. We find that our point estimates are in the 6–9% range for urban workers compared to 7–8% for migrant workers. Returns for men (at 8–9%) are slightly higher than for women (at 6–7%). Thus, private returns to education in urban China in 2009 were not as high as other transition and developing countries, but substantial and have increased over time. Comparing OLS and IV estimates, we also find that the attenuation bias due to measurement error is generally large and more important in the migrant sample compared to the urban sample.

KEYWORDS
Returns to schooling; instrumental variables; rural-to-urban migrants; China

JEL CLASSIFICATION
I21; J61

I. Introduction

The international literature has over the years put forward justifications of why investment in education and the accurate measurement of the return to investing in education has important implications and policy relevance. For example, it guides policies on efficient resource allocation, provides incentives for investing in education by private individuals as well as to what extent the state should subsidize education. Education has also distributional consequences, as returns to education vary across the earnings distribution.

Over the last half century or so, China has undergone large-scale institutional reforms. Until the late 1970s, wages were controlled and based on seniority, rather than productivity; wage differentials by education level were small. The reforms were implemented faster in rural areas. They gave rise to earnings differentials, improved work incentives and spurred growth. As was the case with the Vietnam *doimoi* reforms, urban reforms in China proceeded at a slower pace and were mostly felt after the mid-1990s (Zhang et al. 2005). The other manifestation of a reforming China was the large-scale migration from rural-to-urban areas. It resulted in a large portion of surplus rural labour (estimated at about a quarter of the rural labour force), migrating to urban areas (Lu and Song 2006).

As was the case with other transition economies, as market forces take hold, the expectation is that the profitability of investing in education will increase. For example, in Vietnam during the transition, the return to schooling increased from 3% to 5% during the 1992–1998 period to about 10% post-1998 (Doan and Gibson 2010). There is similar evidence that returns to education have been increasing over the years in China. Better estimates of these returns require dealing with the biases associated with the endogeneity of schooling and unobserved ability. Such studies have been emerging only lately, and there are only a handful of them. In these few studies, estimates of the return to schooling using an instrumental variable (IV) approach vary with the instruments used (although generally found to be significantly higher than the OLS estimates). Questions still remain: have returns reached a level comparable to the world average? Are returns for migrant workers similar to those of urban workers and high enough so more educated workers are incentivized to migrate? Are returns for women

CONTACT Chris Sakellariou acsake@ntu.edu.sg

higher than those of men (as some past studies seem to indicate), or comparable to those of men?

This article uses the 2009 Rural Urban Migration in China (RUMiC), a recently available, rich data set which allows exploration for potentially suitable instruments which can be used for estimating returns to schooling in China. After identifying a set of potential instruments, we conduct a comprehensive validity and relevance testing of different combinations of instruments as well as robustness analysis of our estimates of the return to schooling for rural-to-urban migrants and urban residents in China.

II. Literature review

Past research on returns to education in China varies in focus, examining issues such as the effect of economic reforms on returns, as well as differences in urban, rural and migrant returns. The estimates also vary depending on methodological approach, year of estimation and other factors. There is a general agreement that returns to education have been increasing over time (as was the case with other transition countries), but it is not clear based on existing evidence whether they have reached the developing country average or the average for the region (both at about 10%). We follow with a summary of the literature.

Earlier studies using conventional OLS estimation of Mincerian earning functions found very low returns to schooling. For example, Meng and Kidd (1997) derived estimates of less than 3% for the decade of the 1980s. Fleisher and Wang (2005) used retrospective data for urban residents and found that returns to schooling in China did not begin to increase from the low levels observed at the end of the cultural revolution until nearly 15 years after the initiation of market reforms and approached levels comparable to those in other parts of the world only in the second half of the 1990s; however, they still lagged behind the world average and other transition economies. Low returns to education using the popular Chinese Household Income Project (CHIP) data from the late 1980s were also found by Johnson and Chow (1997) and Liu (1998), among others. Slightly higher returns (at about 5%) were found using the 1995 CHIP data (e.g. see Li 2003). A more recent study by Zhang et al. (2005) used the Mincerian approach and focused on returns to schooling in urban China covering six provinces over an extended period of economic reforms and rising income inequality. They find a dramatic increase in the returns to education, from only 4% per year of schooling in 1988 to 10.2% in 2001. Most of the rise in the returns to education occurred after 1992 and reflected an increase in the wage premium for higher education. De Brauw and Rozelle (2008) looked at returns to schooling in rural China using 2000 data and a different methodological approach. They find that returns are higher than those reported earlier, at about 6.5% and even higher for younger workers and migrants.

Some more recent studies try to address the omitted ability and measurement error biases using mostly IV estimation. Li and Luo (2004) assessed the effect of measurement error and used family background variables such as parental education to control for ability bias. They also used the presence of sons (justified by the Chinese cultural preference for boys) for a smaller sample of young workers as an IV to address ability heterogeneity. They find returns to schooling which are much higher than those from OLS, at about 15%, for young workers in China. Chen and Hamori (2009) used CHNS data from 2004 and 2006. First, they find that OLS estimates are larger than previous studies, at about 7–8%. Using samples of married men and women and spouse's education as instrument, they provide estimates of returns to schooling of 12.5% and 14.5% for married men and women, respectively. The estimate for the return from IV estimation for married women after controlling for sample selection was reported at an unusually high 21%. Heckman and Li (2004) addressed a different problem; that of heterogeneous returns and self-selection into schooling based on such heterogeneous returns. They focus on college attendance and find that for a randomly selected young person from an urban area college attendance leads to a 43% increase in lifetime earnings (nearly 11% annually) in 2000 compared with just 36% (nearly 9% annually) for those who did not attend. They conclude that the return to education has increased substantially in China since the early 1990s.

III. Data

The Survey on RUMiC was generated to study the patterns and effects of migration in China and consists of three parts: the Urban Household Survey, the Rural Household Survey and the Migrant Household Survey.[1] There is particular emphasis on the welfare status of migrants, that is, their jobs, incomes, physical and mental health, their children's education and health, and the extent to which they assimilate into their city communities. The individual-level component covers four areas: (1) household composition; (2) adult education; (3) adult employment and (4) information on children. The household head answered questions covering: (1) social networks; (2) lifecycle events; (3) household income; (4) household assets and (5) housing conditions.

The Rural Household Survey covers nine provinces,[2] and the Urban Migrant Survey covers 15 cities (which are provincial capital cities or other major migrant-receiving cities)[3] in nine provinces or metropolitan areas. The Urban Household Survey was conducted in 19 cities.[4] The distribution of the sample size across the 15 cities is loosely associated with the overall population size of the city. Within each city the sampling frame is defined on the bases of workplaces rather than residence. This is mainly because a sizable proportion of migrant workers in China live in workplace dormitories, construction sites and other workplaces. The sampling design allowed the survey team to estimate the total size of the migrant worker population in each city.

IV. Methodology

Looking for an instrument set

Consider an earnings function with one explanatory variable being potentially endogenous (in our case years of schooling):

$$\ln(W_i) = \beta_0 + \beta_1 Y_i + \beta_2 X_i + \varepsilon_i$$

where W is the wage rate, Y is years of education completed, X is a vector of other controls assumed to be exogenous and ε is the error term. IV estimation entails identifying a set of variables Z (the set of instruments) which is uncorrelated with ε; is correlated with the problematic variable Y; and the variables in Z are not explanatory variables in the original equation. The first-stage (reduced form) estimates are derived from

$$Y_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_i + \mu_i$$

and are used to derive the fitted values of Y_i using OLS.

The bias originates in variable Y being correlated with the disturbance term; similarly, to the extent that variable Y is measured with error, it will be negatively correlated with the disturbance term. The IV estimator (by being a consistent estimator) can avoid the bias that OLS suffers from, when an explanatory variable in a regression is correlated with the regression's disturbance term; however, IV estimation requires both a valid instrument (instrument not itself correlated with the disturbance term and not an independent explanatory variable in the original equation) and an instrument that is not 'too weak' (i.e. it is sufficiently correlated with the endogenous explanatory variable). In what follows, we outline the main considerations guiding our methodological approach[5] in estimating the return to schooling for rural-to-urban migrants and urban residents in China using IV estimation.

Although having as many instruments as endogenous regressors is sufficient for identification, it is desirable to have more suitable instruments than required. This is because the two-stage least-squares estimator has larger SEs compared to the OLS estimator (generally, several times higher); hence, a larger number of over-identifying restrictions tends to result in a higher R^2 in the first stage, which results in smaller standard errors. Furthermore, one can exploit such over-identification to test the validity of individual instruments. It is also desirable that not all instruments in the set are based on a common rationale; tests of over-identifying restrictions

[1]RUMiC was initiated by a group of researchers at the Australian National University, the University of Queensland and the Beijing Normal University and was supported by the Institute for the Study of Labor (IZA), which provides the Scientific Use Files. The financial support for RUMiC was obtained from the Australian Research Council, the Australian Agency for International Development (AusAID), the Ford Foundation, IZA and the Chinese Foundation of Social Sciences.

[2]These are Anhui, Chongqing, Guangdong, Hebei, Henan, Hubei, Jiangsu, Sichuan and Zhejiang.

[3]These are Bengbu, Chengdu, Chongqing, Dongguan, Guangzhou, Hefei, Hangzhou, Luoyang, Nanjing, Ningbo, Shanghai, Shenzen, Wuhan, Wuxi and Zhengzhou.

[4]It includes the following additional cities to the Urban Migrant Survey: Anyang, Jiande, Leshan and Mianyang.

[5]Murray (2006) provides an excellent discussion for what one should consider when assessing validity and strength of instruments.

test instrument validity while assuming that there are enough valid instruments for at least exact identification. Since satisfying the over-identifying restrictions does not mean that *all* the instruments are necessarily valid, it helps to have a mix of instruments of varying rationales. Yet another use of having several potentially valid instruments is that one can use different combination of subsets of these instruments; do coefficients estimates vary widely or remain approximately the same? The second outcome would enhance the credibility of the instrument set.

Another concern is to what extent the instrument set qualifies as 'strong' (exhibit sufficient correlation with the endogenous regressor). The relevance of this is because when the instrument set is weak coefficient estimates will be biased (though consistent) with finite samples; the weaker the instruments, the larger the bias. Furthermore, with weak instruments, SEs of estimates are biased downwards, resulting in misleading confidence intervals. The extent of the coefficient bias depends on the number of instruments, the strength of instruments (the R^2 of the reduced form regression) and the sample size. It increases with the number of instruments (hence a trade-off: more instruments, higher R^2 along with an increase in the bias) and it decreases with the R^2 of the first-stage regression, as well as with a higher sample size. As a rule of thumb, if the product of sample size times the R^2 exceeds the number of instruments, the IV estimates will be less biased compared to the OLS estimates. The appropriate test for evaluating whether the instrument set is weak is the Stock–Yogo (2005) test.[6] Similarly, the Stock–Yogo test can be used to evaluate the hypothesis that the true significance level of the endogenous regressor is smaller than say 10%, for a stated significance level of 5%. Finally, having a strong set of instruments is important because even if there are instruments which are not clearly valid (i.e. 'almost valid'), if the instrument set is strong the bias is likely to be limited.

One could also state preference for instruments which apply to the entire sample, as opposed to a restricted/selected sample. For example, in estimating the return to schooling, it is common to use restricted samples if the chosen instrument is, say, spouse's education (applies only to those who are married) or parents' education (it applies only to children of the head of household, if the data do not contain information on education of parents for all individuals).

The instrument set

After extensive exploration, the instrument set used is based on the following information available in the RUMiC 2009 surveys: (1) sibling composition: being firstborn versus later-born, along with the interaction with age cohort and/or number of siblings; and for women, presence of a male sibling interacted with age and with number of siblings (available for the urban subsample only); (2) age of school entry: age 6 versus age 7 (available only in the Migrant survey), along with its interaction with age cohort; and (3) spouse's years of schooling completed, which is used only for robustness checks of our estimates. Below we discuss each of the instruments used.

Sibling composition

There is an emerging literature exploring the degree to which number and gender composition of siblings and birth order affect a child's subsequent educational attainment. This is based on the theory suggesting a trade-off between child quantity and 'quality'. One can argue that siblings are not necessarily expected to receive equal shares of resources and attention by parents in education acquisition; furthermore, a larger family size might adversely affect the production of child quality within a family.

Booth and Kee (2005) used British Household Panel Survey data to explore the degree to which family size and birth order affect a child's subsequent educational attainment. They find that siblings are not assigned equal shares in the family's educational resources; instead, the shares are decreasing with birth order. They also found that given the birth order effect the family size effect does not vanish once they control for birth order. Fergusson, Horwood, and Boden (2006) used New Zealand data and nested models to control for the confounding effects of family size on birth order and found that

[6]In the Stock–Yogo test, the hypothesis is that the IV bias is less than some fraction (say 5%) of the OLS bias.

birth order effects on educational attainment were not disguised by family size effects. A statistically significant association remained between being later-born and a lower likelihood of obtaining educational qualifications. They concluded that the intra-family dynamics initiated by birth order may have a lasting effect on the individual in terms of later educational and achievement outcomes.

Bagger et al. (2013) used an empirical strategy that identifies the effect of family size on the intra-household distribution of human capital separately from the effect that birth order may have on a child's education using Danish data; their results suggest that both birth order and family size affect years of education, confirming the presence of a quantity–quality trade-off. They found that birth order has a strong negative effect on a child's education, consistent with existing empirical studies. Overall, they provided evidence supporting the existence of a trade-off between quality and quantity of children. In a recent paper, Bu (2014) used sibling data from the British Household Panel Survey and found that firstborn children enjoy a distinct advantage over their later-born counterparts in terms of educational attainment. In particular, she found that firstborn children have higher aspirations, and that these aspirations play a significant role in determining later levels of attainment.

In the Chinese context, Qian (2009) exploited plausibly exogenous changes in family size caused by relaxations in China's One Child Policy to estimate the causal effect of family size on school enrolment of the first child. The results show that for one-child families an additional child significantly increased school enrolment of firstborn children by approximately 16 percentage-points. She also found that the 1 son–2 children relaxation increased family size for girls born in areas affected by the relaxation.

The one-child policy was introduced in 1979. It was subject to exceptions: rural families can have a second child if the first child is a girl or is disabled, and ethnic minorities are exempt. Beginning in 1987, official policy granted local officials the flexibility to make exceptions and allow second children in the case of 'practical difficulties' or when both parents are single children[7]; some provinces had other exemptions worked into their policies as well (e.g. Sichuan allowed exemptions for couples of certain backgrounds). After the introduction of the one-child policy, the fertility rate in China fell from 2.63 births per woman in 1980 (already a sharp reduction from more than five births per woman in the early 1970s) to 1.61 in 2009 (World Development Indicators 2009). However, it is understood that the policy was probably only partially responsible for the reduction in the total fertility rate (Hesketh, Lu, and Xing 2005). Chart 1 in the Appendix depicts overtime declines in the average number of siblings for adults born from 1964 to 1987 using the urban and rural-to-urban migrant files in the 2009 RUMiC surveys.

On theoretical grounds, the validity of an instrument based on birth order requires that birth order is unrelated to unobserved ability. There is an extensive literature spanning several decades (mostly from Psychology) investigating the relationship between intelligence and birth order. Taking Belmont and Marolla (1973) as a starting point, the authors provided an empirical compilation of Raven Progressive Matrices scores from a cross section of almost 400 000 Dutch men of different birth orders. When the IQ scores were disaggregated by levels of birth order and family size, a systematic pattern seemed to emerge, which suggested declining intelligence with increasing birth order and family size. They cautioned, though, that the differences, although highly systematic, were very small. Several such papers subsequently appeared leading to the 'confluence' hypothesis (e.g. Zajonc and Markus 1975; Zajonc 1976; Zajonc and Mullally 1997).

However, more recent papers provide strong evidence that the 'negative birth order' phenomenon is likely a methodological illusion. In particular, Rodgers et al. (2000) compared the patterns from past cross-sectional data to those from the few within-family studies and found that they are entirely different; that is, the negative birth order effect disappeared when the IQ measures of actual

[7]This policy was implemented at different times in different provinces. Zhejiang, the first province to adopt such a policy, introduced it in 1985. By the end of the last century, 27 provinces and municipalities have passed it, but it achieved full coverage in China only in 2011 when Henan government finally agreed to the implementation of such a policy. In order to address the demographic challenges such as ageing population and shrinking labour force, the one-child policy was relaxed further in the Third Plenary Session of the 18th CPC Central Committee in 2013, whereby families with one parent being the only child have gradually been permitted to have two children (the time of implementation is subject to the revision of the regulation of the local government).

siblings were compared to one another. Whichman, Rodgers, and McCallum (2006) observed that if mean IQ scores decline across birth order the cause of those declines may lie either within the family or outside of the family. He questioned the practice by researchers that have consistently interpreted those causes to lie within the family and have built within-family models to explain the declines. They used National Longitudinal Survey of Youth data and compared siblings to one another at fixed ages and found conclusive evidence that the fundamental cause of presumed birth order effect lies between, not within, families. Thus, using a different methodology, they come to the same conclusion as most other recent studies, that the sources of the often found birth order–intelligence relationship appear to lie outside the family.

When estimating returns to schooling for women, another candidate instrument based on sibling composition is presence of a male sibling. Conditional on family size (or number of siblings), daughters may obtain less education than their male siblings, if families prefer sons and allocate more resources to sons compared to daughters (Chen, Chen, and Liu 2009). The argument for using sibling gender as an instrument is that the gender composition of siblings can significantly affect educational attainment but is not related to inherent ability. Li and Luo (2004) used presence of brothers to estimate returns to schooling for Chinese women; they argued that in the Chinese context, families prefer boys, as boys carry on the family name and assume the responsibility of care for their elderly parents. Presence of boys can also affect fertility decisions as it increases the probability of not having another child upon the birth of a boy.

Age of school entry

Generally, school entry age varies due to the use of a single school cut-off date. Evidence on the effect of age of school entry on educational attainment and school performance (such as probability of failing a grade, highest degree attained, etc.) is less than conclusive. Fertig and Kluve (2005) used a data set of children entering school during the 1970s in West and East Germany and alternative estimation approaches; when a linear probability model or a matching approach was used, they found a qualitatively negative relation between the age at school entry and educational outcomes both in terms of schooling degree and probability of having to repeat a grade (i.e. an older age at school entry is associated with a higher probability to repeat a class, a lower probability to receive a high schooling degree in West Germany and a higher probability to drop out of school). However, when an IV approach was followed (using a cut-off date rule and the corresponding age at school entry according to the regulation to instrument the actual age at school entry), estimates suggest there is no effect of age at school entry on educational performance. The authors suggest that it is likely that these findings could be driven by unobserved heterogeneity, that is, those individuals who entered late did so because they were conjectured (by their parents or elementary school teachers) to display low educational performance.

Evidence on the effect of age of school entry on school performance tends to agree (at least based on research on northern European countries) that entering school at 7 is associated with a better school performance. Puhani and Weber (2005) used three German data sets and IV estimation and found robust and significant positive effects on educational attainment for pupils who enter school at 7 instead of 6 years of age; in particular, they found that test scores at the end of primary school increase by about 0.42 SDs and years of secondary schooling increase by almost half a year. Similar findings have been reported by Bedard and Dhuey (2006) for Sweden and Strøm (2004) for Norway. However, these differ from those of Angrist and Krueger (1992) and Mayer and Knutson (1999) for the United States, where either no or negative effects for late school entry are reported.

In China, the two-semester school year usually begins on 1 September and sixteenth of the first month of Chinese lunar year, with a summer vacation in July and August and a winter vacation around the Chinese spring festival; however, over the years there were periods with a spring enrolment (April). Chart 2 in the Appendix shows the proportion of individuals in the file of rural-to-urban migrants (age of school entry information is not available for urban residents), who entered school at age 7 by month of birth; it can be seen that this proportion spikes for those born right after the months of September and April.[8]

According to the provisions of the Compulsory Education Law of the People's Republic of China, the 6 years of primary education start at age 6 (fully 6 years old) or 7; children usually entered primary school at 7 years of age, although over the years the proportion of children who entered school at age 6 in China has been increasing. Chart 3 shows the variation in the proportion of those who entered school at age 6 by age cohort: the more recent the cohort, the higher the proportion who entered school at age 6 (as opposed to age 7).

Age of school entry as an instrument will be valid if it is uncorrelated with unobserved ability. The school cut-off date for entry can be taken as exogenous; however, to the extent parents can manipulate age of entry to primary school for their children based on perceived ability, unobserved heterogeneity could be an issue. On the other hand, the finding that over time the proportion of children entering school earlier has increased seems to suggest that increasingly parents believe that their children enjoy an advantage (head start) by starting school earlier. To reduce potential unobserved heterogeneity, we eliminated from the sample the small proportion of observations with reported age of entry lower than 6 years and greater than 7 years.

Spouse's education

We are using this instrument to evaluate the robustness of the main findings. Spouse's education is suggested as a possible valid instrument by Trostel, Walker, and Woolley (2002), who explored the independence of wife's education from husband's earnings and its interaction with husband's education. These studies rely on the assortative nature of marriage, as married couples share common interests and behavioural traits, and they usually share a common level of schooling (Pencavel 1998). Trostel, Walker, and Woolley (2002) obtained estimates using spouse's education to instrument for schooling that are over 20% higher than the corresponding OLS estimates, suggesting that conventional OLS estimates might be biased downwards. Arabsheibani and Mussurov (2007) and Lall and Sakellariou (2010) also find that spouse's education is a valid instrument and that the conventional OLS estimates, which do not control for endogeneity bias, might underestimate the true return to education. In the Chinese context, Chen and Hamori (2009) used spouse's education as an instrument and found higher returns to schooling from two-stage least squares compared to OLS, especially for women.

V. Estimation and discussion

Estimation results

The estimation sample used is for workers 22–54 years of age with positive earnings. In the earnings function specification, the dependent variable is the logarithm of hourly wage derived using the information on monthly earnings from the primary job (including bonus and payments in kind) and hours worked. Education is measured as the years of schooling completed (excluding skipping or failing a grade). The data include information on the actual years of tenure in the current job; so, instead of using years of potential experience and its square (or age and its square) we are able to use years of tenure and years of tenure squared; we also included years of other (potential) experience and its square. Other characteristics controlled for are marital status and size of firm. The results obtained are OLS estimates from Mincerian (Mincer 1974) earnings functions, selectivity corrected estimates using Heckman correction (for female workers) and IV estimates. Information on the instruments used is available for all members of the household, with some missing values; information on sibling composition is available for 91% of persons for the migrant sample and 78% of persons in the urban sample; information on age of school entry is available for 73% of persons after excluding those who entered school before age 6 or later than age 7 (available only in the migrant survey).

[8]The figure is drawn for the subsample of those born from 1964 to 1987, who started school at the age of either 6 or 7 between 1970 and 1993; during this period, there were indeed some years with a spring enrolment.

Urban workers

The instrument set for estimating returns to schooling for urban workers is based on sibling composition only since information on age of school entry is not available in the urban survey. Table 1 presents the results for men. The instruments are being a firstborn child and its interaction with age. The OLS estimate of the return to schooling is just over 7%. The specification included the firstborn dummy in the controls; its coefficient is insignificant, suggesting that there is no independent effect of being firstborn on male earnings. The estimate of the return to schooling from IV estimation is 2 percentage-points higher, at about 9%. The instrument set is strong, with an F-value of about 80, which exceeds the critical values for the 5% maximal IV relative bias and the 10% maximal IV size in the Stock–Yogo weak identification tests. Inspection of the validity tests shows that there is no indication that the instrument set is invalid. Given that the difference between the OLS and IV estimates is not large, the endogeneity test for years of schooling does not reject the hypothesis that schooling is exogenous (p-value of 0.38).

The marginal return to an additional year of tenure (partial derivative evaluated at mean years of tenure) is concave, with earnings increasing by about 2% per additional year. The return to other potential experience is also concave, but its effect on earnings is much smaller. The premium associated with being a married male is about 25%. Finally, there is a significant premium associated with working in larger firms. Similar effects of other covariates apply also to the rest of the results.

The results for women (see Table 2) generally point to lower returns to schooling compared to those for men. The instruments used are presence of a male sibling interacted with age and with number of siblings (instrument set 1) and those two along with being firstborn interacted with age (instrument set 2). The point estimate from OLS is 5.7% compared to 7.2% for males (statistically significant based on 5% confidence intervals). Selectivity corrected estimates are of similar magnitude (at 5.4%), with a positive and significant *lambda*. The IV and OLS estimates are of similar magnitude, at about 5.5% for the OLS estimates and 6–6.5% for the IV estimates. The instrument sets are strong and there are no major issues with the validity of the instrument set (although the Sargan statistic p-value is not very high for the second instrument set). The endogeneity test p-value reflects the similarity of OLS and IV estimates of the return to schooling.

Robustness checks were conducted, using a sample of married urban workers and spouses' years of education as the main instrument (see Tables A1 and A2). The other instrument sets are for men, a combination of spouses years of schooling, being firstborn and its interaction with age (column 3) and being firstborn and its

Table 1. Male urban workers.

	OLS[a]	IV
Years of schooling	**0.072**	**0.091**
	(0.006)	(0.023)
Tenure	**0.034**	**0.037**
	(0.006)	(0.006)
Tenure squared	**−0.0008**	**−0.0008**
	(0.0001)	(0.0001)
Other experience	**0.012**	**0.019**
	(0.005)	(0.009)
Other experience squared	**−0.0006**	**−0.0007**
	(0.0001)	(0.0002)
Married	**0.258**	**0.232**
	(0.050)	(0.051)
Firm size: 6–20	**0.212**	**0.194**
	(0.058)	(0.054)
Firm size: 21–100	**0.322**	**0.299**
	(0.054)	(0.052)
Firm size: >100	**0.324**	**0.303**
	(0.053)	(0.050)
Firstborn	−0.010	–
	(0.026)	
Constant	**0.964**	0.669
	(0.106)	(0.343)
First stage		
Shea partial R^2/partial R^2		0.046
F-Value [p-value]		80.24
		[0.000]
Second stage		
Over-identification test: all instruments: Sargan statistic [p-value]		0.188
		(0.665]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for		
- Firstborn		–
- Firstborn interacted with age		–
Weak identification test		
Cragg–Donald Wald statistic (F-statistic)		80.24
10% maximal IV size		19.93
Redundancy test (χ^2 *[p-value]* for		
- Firstborn		140.7
		[0.000]
- Firstborn interacted with age		152.4
		[0.000]
Endogeneity test for years of schooling		
χ^2 [p-value]		0.756
		[0.385]
R^2	0.191	
N	3316	3316

Notes: Significant at the 5% level or less in bold.
[a]Robust SEs in parentheses; OLS results using the same sample as for IV estimation.

Table 2. Female urban workers.

	OLS[a]	Heckman[b]	IV	IV
Years of schooling	**0.057**	**0.054**	**0.050**	**0.063**
	(0.006)	(0.006)	(0.015)	(0.013)
Tenure	**0.026**	**0.024**	**0.025**	**0.027**
	(0.006)	(0.006)	(0.006)	(0.006)
Tenure squared	**−0.0006**	**−0.0007**	**−0.0006**	**−0.0006**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Other experience	−0.007	−0.008	−0.010	−0.004
	(0.005)	(0.006)	(0.008)	(0.007)
Other experience squared	−0.0001	−0.0003	−0.0000	−0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Married	0.058	0.030	0.067	0.054
	(0.043)	(0.067)	(0.044)	(0.044)
Firm size: 6–20	**0.116**	**0.129**	**0.130**	**0.105**
	(0.053)	(0.050)	(0.053)	(0.051)
Firm size: 21–100	**0.317**	**0.340**	**0.328**	**0.310**
	(0.053)	(0.048)	(0.048)	(0.047)
Firm size: >100	**0.314**	**0.321**	**0.328**	**0.305**
	(0.051)	(0.047)	(0.049)	(0.047)
Has brother/s	0.041	0.041	0.043	0.043
	(0.034)	(0.035)	(0.034)	(0.034)
Number of siblings	−0.017	−0.015	−0.012	−0.018
	(0.011)	(0.011)	(0.010)	(0.011)
Firstborn	−0.046	−0.053	–	−0.047
	(0.031)	(0.032)		(0.031)
Constant	**1.28**	**1.27**	**1.36**	**1.18**
	(0.116)	(0.135)	(0.240)	(0.199)
Lambda		**0.336**		
		(0.101)		
First stage				
Shea partial R^2/partial R^2			0.119	0.179
F-Value [p-value]			164.5 [0.000]	176.4 [0.000]
Second stage				
Over-identification test: all instruments				
Sargan statistic [p-value]			0.698 [0.439]	0.257 [0.277]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for				
- Has brother/s interacted with age			–	2.08 [0.149]
- Has brother/s interacted with # of siblings			–	0.095 [0.758]
- Firstborn interacted with age			–	1.84 [0.175]
Weak identification test				
Cragg–Donald Wald statistic (F-statistic)			164.5	176.4
Stock–Yogo weak ID test critical values				
5% maximal IV relative bias			–	13.91
10% maximal IV size			19.93	22.30
Redundancy test (χ^2 [p-value] for				
- Has brother/s interacted with age			290.9 [0.000]	391.3 [0.000]
- Has brother/s interacted with # of siblings			35.82 [0.000]	11.28 [0.001]
- Firstborn interacted with age			–	166.9 [0.000]
Endogeneity test for years of schooling				
χ^2 [p-value]			0.285 [0.593]	0.036 [0.551]
R^2	0.209			
Wald χ^2 [p-value]		612.7		
N	2451	3119 (974 censored)	2451	2451

Notes: Significant at the 5% level or less in bold.
[a]Robust SEs in parentheses; OLS results using the same sample as for IV estimation.
[b]Independent variables in selection equation: age, age squared, marital status, number of children, majority ethnic group.

interaction with age without spouse's years of schooling (column 4). For women, a combination of spouse's years of schooling, presence of male sibling interacted with age and number of siblings (column 3), only presence of male sibling interacted with age and number of siblings (column 4) and the same two along with being firstborn interacted with age (column 5).

OLS and IV estimates for married male urban workers are very close to those in the main results, with the exception for those in column 4, where the point estimate is higher, at about 13% compared to 9% in Table 1 using the same combination of instruments; this difference likely reflects the use of a selected sample (that of married men). Similar robustness checks using married female urban workers show that the OLS estimates as well as the IV estimates in columns 4 and 5 are similar to those in the main results and somewhat higher when the spouse's years of schooling instrument is involved (columns 2 and 3). Instrument sets are strong and tests validate the instruments in every specification.

Migrant workers

Table 3 presents the results for male migrant workers. Column 1 gives the OLS estimates using the same sample as for the IV estimation, while columns 2–4 give the IV estimates for different combinations of instruments. The OLS estimate of the return to schooling is 4.5%. The independent effect of being firstborn and having entered school before the age of 7 on earnings is insignificant. The full set of instruments available is based on being firstborn along with its interaction with age, and having entered school before the age of 7 along with its interaction with age; the interaction with age is used because of the overtime increasing proportion of schoolchildren entering school at age 6 rather than age 7 (see Chart 3).

Table 3. Male migrant workers.

	OLS[a]	IV	IV	IV
Years of schooling	**0.045**	**0.080**	**0.080**	**0.084**
	(0.009)	(0.034)	(0.035)	(0.038)
Tenure	**0.034**	**0.038**	**0.038**	**0.038**
	(0.007)	(0.010)	(0.010)	(0.010)
Tenure squared	**−0.0011**	**−0.0011**	**−0.0011**	**−0.0011**
	(0.0003)	(0.0004)	(0.0004)	(0.0004)
Other experience	**−0.014**	−0.003	−0.003	−0.002
	(0.007)	(0.012)	(0.012)	(0.013)
Other experience squared	0.0001	0.0000	0.0000	0.0000
	(0.0002)	(0.000)	(0.000)	(0.000)
Married	**0.090**	0.067	0.067	0.064
	(0.036)	(0.048)	(0.048)	(0.049)
Firm size: 6–20	**0.168**	**0.167**	**0.167**	**0.147**
	(0.052)	(0.050)	(0.050)	(0.050)
Firm size: 21–100	**0.296**	**0.284**	**0.284**	**0.283**
	(0.052)	(0.050)	(0.050)	(0.050)
Firm size: >100	**0.368**	**0.351**	**0.351**	**0.349**
	(0.048)	(0.045)	(0.045)	(0.046)
Firstborn	0.010	–	–	–
	(0.033)			
Started school before age 7	−0.001	–	–	–
	(0.032)			
Constant	**1.25**	**0.838**	0.839	0.790
	(0.114)	(0.412)	(0.425)	(0.459)
First stage				
Shea partial R^2/partial R^2		0.053	0.048	0.043
F-Value [*p*-value]		19.30 [0.000]	23.47 [0.000]	31.08 [0.000]
Second stage				
Over-identification test: all instruments				
Sargan statistic [*p*-value]		0.075 [0.995]	0.070 [0.791]	0.001 [0.978]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for				
- Firstborn		0.021 [0.884]	–	–
- Firstborn interacted with age		0.038 [0.846]	0.051 [0.821]	–
- Started school at age 6 (versus at age 7)		0.005 [0.945]	0.005 [0.945]	–
- Started school at age 6 interacted with age		0.018 [0.895]	0.018 [0.895]	–
Weak identification test				
Cragg–Donald Wald statistic (*F*-statistic)		19.30	23.47	31.08
Stock–Yogo weak ID test critical values				
5% maximal IV relative bias		16.85	13.91	19.93
10% maximal IV size		–	22.30	–
15% maximal IV size		13.96	–	–
Redundancy test (χ^2 [p-value]) for				
- Firstborn		41.37 [0.000]	–	42.12 [0.000]
- Firstborn interacted with age		52.63 [0.000]	52.63 [0.000]	52.73 [0.000]
- Started school at age 6 (versus at age 7)		6.55 [0.010]	6.55 [0.010]	–
- Started school at age 6 interacted with age		10.09 [0.001]	10.09 [0.001]	–
Endogeneity test for years of schooling				
χ^2 [*p*-value]		1.13 [0.287]	1.03 [0.310]	1.13 [0.287]
R^2	0.141			
N	1405	1405	1405	1405

Notes: Significant at the 5% level or less in bold.
[a]Robust SEs in parentheses; using the same sample as in IV estimation.

The estimate of the return to schooling is now higher, at 8%. The p-values from testing the over-identifying restrictions and the exogeneity/orthogonality of suspect instruments are all very high, suggestive of the validity of the combination of instruments. Based on the partial R^2 (0.053) and the F-value (19.3), the instrument set is fairly strong. Columns 3 and 4 give the IV results using alternative instrument sets: firstborn interaction with age and the two instruments based on age of school entry in column 3 and the two instruments based on age of school entry only in column 4. The estimates of the return to schooling are essentially the same, at 8–8.5%, and so are the conclusions from testing over-identifying restrictions.

Table 4 presents the results for female rural-to-urban migrants. While point estimates of OLS and IV-based returns are slightly lower compared to those for men, based on confidence intervals they are not statistically different. The OLS and selectivity

Table 4. Female migrant workers.

	OLS[a]	Heckman[b]	IV	IV
Years of schooling	**0.036**	**0.033**	**0.074**	0.072
	(0.008)	(0.009)	(0.032)	(0.040)
Tenure	**0.049**	**0.044**	**0.057**	**0.054**
	(0.012)	(0.012)	(0.013)	(0.014)
Tenure squared	**−0.0021**	**−0.0018**	**−0.0023**	**−0.0022**
	(0.0008)	(0.0007)	(0.0007)	(0.0007)
Other experience	0.001	-0.000	0.008	0.011
	(0.007)	(0.008)	(0.012)	(0.014)
Other experience squared	−0.0003	−0.0002	−0.0003	−0.0004
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Married	−0.053	−0.039	−0.020	−0.044
	(0.044)	(0.051)	(0.053)	(0.051)
Firm size: 6–20	**0.127**	**0.123**	0.111	**0.112**
	(0.050)	(0.051)	(0.058)	(0.056)
Firm size: 21–100	**0.242**	**0.243**	**0.273**	**0.220**
	(0.053)	(0.053)	(0.058)	(0.058)
Firm size: >100	**0.225**	**0.218**	**0.243**	**0.207**
	(0.046)	(0.048)	(0.052)	(0.051)
Firstborn	**0.081**	**0.094**	0.039	–
	(0.040)	(0.038)	(0.042)	
Started school before age 7	−0.019	−0.012	–	–
	(0.041)	0.044		
Constant	**1.20**	**1.29**	0.558	0.790
	(0.115)	(0.147)	(0.359)	(0.505)
Lambda		−0.283		
		(0.332)		
First stage				
Shea partial R^2/partial R^2			0.075	0.045
F-Value [p-value]			23.13 [0.000]	22.05 [0.000]
Second stage				
Over-identification test: all instruments				
Sargan statistic [p-value]			1.00 [0.606]	1.38 [0.239]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for				
- Started school at age 6 (versus age 7)			0.578 [0.447]	–
- Started school at age 6 interacted with age			0.357 [0.550]	–
- Firstborn interacted with age			0.051 [9.821]	–
Weak identification test			23.13	22.05
Cragg–Donald Wald statistic (F-statistic)				
Stock–Yogo weak ID test critical values				
5% maximal IV relative bias			13.91	–
10% maximal IV size			22.30	19.93
Redundancy test (χ^2 [p-value] for				
- Started school at age 6 (versus age 7)			8.47 [0.004]	13.47 [0.000]
- Started school at 6 interacted with age			15.94 [0.000]	24.01 [0.000]
- Firstborn interacted with age			32.89 [0.000]	–
Endogeneity test for years of schooling				
χ^2 [p-value]			1.60 [0.206]	0.78 [0.377]
R^2	0.133			
Wald χ^2 [p-value]		119.5		
N	945	1.047 (censored 126)	945	945

Notes: Significant at the 5% level or less in bold.
[a]Robust SEs in parentheses; OLS results using the same sample as for IV estimation.
[b]Independent variables in selection equation: age, age squared, marital status, number of children, majority ethnic group.

corrected estimates of the return to schooling are low, at about 3.5%, and the coefficient of the inverse Mills ratio is insignificant; this is expected given that only a small proportion of observations are censored (i.e. the overwhelming proportion of female migrants are employed). The IV estimates in column 3 are from a combination of instruments based on both rationales, while those in column 4 are based on age of school entry only. Both instrument combinations are associated with a nearly identical estimate of the return to schooling, at 7–7.5%. The instrument set in column 3 is strong and there is no indication of any problem associated with validity of instruments; the instrument set in column 4 is fairly strong, while the test of over-identifying restrictions is associated with a somewhat low p-value (0.24).

Finally, we have used different age cohorts as the estimation sample and found that the return to schooling is higher for younger cohorts; for example, using workers aged 22–44, the estimates are 1–2 percentage-points higher. This is likely because with younger cohorts a larger proportion obtained their education and entered the labour market during the post-transition period. This may explain the high estimates of the return to schooling for young Chinese workers (about 10% from OLS and 17% or higher from IV estimation) derived in Li and Luo (2004).

Summarizing, the main findings show that, as generally found in empirical research on returns to education, IV estimates exceed the corresponding OLS estimates. The difference in estimates is particularly large for migrants, suggesting (perhaps in accordance with intuition) that the attenuation bias due to measurement error is more important in the migrant sample compared to the urban sample. Our estimates, although not very high, suggest that there are substantial and overtime increasing incentives to invest in human capital for rural migrants and government funding for education in emigration regions.

On the other hand, the OLS and IV estimates differ much less for urban workers (especially urban female workers). We also find that the return to schooling for male urban workers is about 1–2% higher than for female urban workers based on OLS estimates and 2–3% higher based on IV estimates. Our estimates differ from some previous studies such as Li (2003) and Chen and Hamori (2009); for example, Chen and Hamori (2009) find that the OLS estimates are higher for men while the IV estimates are higher for women. Size wise, returns to education in urban China in 2009 are not as high as those for other transition countries (e.g. Vietnam), as well as to worldwide and developing country averages (of about 10%), but have been increasing over the years.

Implications

Deriving reliable estimates of the return to education can be useful to policymaking. In doing so, one needs to account for endogeneity and measurement error biases. IV estimation deals with such biases. Over the last decade or so, a handful of studies using this methodology have appeared for China; however, estimates of the return to schooling differ substantially despite the common finding that IV-based return estimates are substantially higher compared to those from OLS regression. There are differences in data used (sometimes nationally representative, others regional) as well as instruments used. Using instruments such as parental education and spouse's education would result in a selected sample. On the other hand, use of various single instruments can lead to drastically different estimates of the return to schooling, as the size of estimates can depend on the instrument used (as opposed to a combination of potentially valid instruments). Some past return estimates based on IV estimation for China are unusually large (in the order of 15–20%).

We have used a new nationally representative data set particularly suitable for estimating returns to education for rural-to-urban migrants and derived IV-based return estimates for urban and migrant workers using combinations of instruments, paying particular attention to validity and relevance testing. Our estimates, while confirming that returns to education in China have risen over the transition, are generally lower compared to other IV-based estimates in the empirical literature. We also find that IV estimates are generally higher than OLS estimates, but such differences are clearly evident only for migrants. OLS and IV-based return estimates for male urban workers are quite close, at 7% and 9%, respectively, while the corresponding return estimates for female urban workers are of similar magnitude at about 6%. On the other hand, the IV-based estimates for migrants are about double those from OLS (8–8.5% versus 4.5% for men and 7–7.5% versus 3.5% for women), suggesting that the downward bias associated with measurement error is particularly large in the migrant sample.

We find that returns for rural-to-urban migrants are substantial and comparable to those for urban workers. Thus, despite the *hukou* system the incentive for migration of skilled workers is present; however, the *hukou* system could be fine-tuned (if not abolished as suggested in Lu and Song 2006) by exempting prospective migrants who undergo pre-migration skill upgrading training based on skills deemed important in enhancing productivity in urban labour markets. It is also important to provide further incentives for human capital and skill acquisition in rural areas in order to reduce the cost of investing in human capital (in addition to exemptions of school fees and related expenses for compulsory education), given that returns to education in rural areas are lower compared to urban areas in China.

VI. Conclusion

This article contributes to the returns to education in China and IV estimation literature by utilizing a new rich data set, new instrument combinations and a comprehensive evaluation of estimates against the challenges faced by researchers in justifying IV estimates. We find that the private return to an additional year of schooling in China, after accounting for endogeneity of schooling and measurement error, after rising during China's transition and transformation of labour markets, are still somewhat lower than the averages for emerging economies and other developing countries. Male returns are of similar magnitude to female returns for migrant workers, while our estimates for urban workers are higher for men. The IV estimates are higher compared to the OLS estimates and much more so for migrant workers; thus, measurement error of the schooling variable results in a substantial downward bias of the OLS estimates, especially for rural-to-urban migrant estimates. Returns to schooling for men are of similar magnitude for urban and migrant workers, while for women the IV estimates for migrant female workers are slightly higher than those for urban female workers, but the difference is not statistically significant. The results suggest that returns to education in China have risen over time to levels not far below those in other transition countries, at least for men. The returns for migrant workers at 7–8% are substantial and provide incentives to invest in human capital for rural migrants and justifying government funding for education in emigration regions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Angrist, J., and A. Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association* 87: 328–336. doi:10.1080/01621459.1992.10475212.

Arabsheibani, R., and A. Mussurov. 2007. "Returns to Schooling in Kazakhstan." *Economics of Transition* 15 (2): 341–364. doi:10.1111/ecot.2007.15.issue-2.

Bagger, J., J. Birthenall, H. Mansour, and S. Urzua. 2013. "Education, Birth Order and Family Size." *IZA Discussion Paper No. 7454*, Bonn, Germany: Institute for the Study of Labour.

Bedard, K., and E. Dhuey. 2006. "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects." *Quarterly Journal of Economics* 121: 1437–1472.

Belmont, L., and F. A. Marolla. 1973. "Birth Order, Family Size, and Intelligence: A Study of a Total Population of 19-Year-Old Men Born in the Netherlands Is Presented." *Science* 182: 1096–1101. doi:10.1126/science.182.4117.1096.

Booth, A., and H. J. Kee. 2005. "Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment." *Discussion Paper No 506*, Canberra, Australia: Centre for Economic Policy Research, Australian National University.

Bu, F. 2014. "Sibling Configurations, Educational Aspiration and Attainment." *Discussion Paper No. 2014-11*. Essex: Institute for Social and Economic Research, University of Essex.

Chen, G., and S. Hamori. 2009. "Economic Returns to Schooling in Urban China: OLS and the Instrumental Variables Approach." *China Economic Review* 20: 143–152. doi:10.1016/j.chieco.2009.01.003.

Chen, Y. C., S. H. Chen, and J. T. Liu. 2009. "Separate Effects of Sibling Gender and Family Size on Educational Achievements – Methods and First Evidence from Population Birth Registry." *Discussion Paper 2008 – 3.*

London: Department of Economics, Royal Holloway College, University of London.

De Brauw, A., and S. Rozelle. 2008. "Reconciling the Returns to Education in Off-Farm Wage Employment in Rural China." *Review of Development Economics* 12 (1): 57–71.

Doan, D., and J. Gibson. 2010. "Return to Schooling in Vietnam during Economic Transition: Did the Return Reach Its Peak?" *Paper No. 24984*, Munich Personal RePEc Archives (October), Munich.

Fergusson, D. M., J. Horwood, and J. Boden. 2006. "Birth Order and Educational Achievement in Adolescence and Young Adulthood." *Australian Journal of Education* 50 (2): 122–139. doi:10.1177/000494410605000203.

Fertig, M., and J. Kluve. 2005. "The Effect of Age at School Entry on Educational Attainment in Germany." *Discussion Paper No. 27*, Essen, Germany: RWI.

Fleisher, B. M., and X. Wang. 2005. "Returns to Schooling in China Under Planning and Reform." *Journal of Comparative Economics* 33: 265–277. doi:10.1016/j.jce.2005.03.006.

Heckman, J. J., and X. Li. 2004. "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." *Pacific Economic Review* 9 (3): 155–171. doi:10.1111/per.2004.9.issue-3.

Hesketh, T., L. Lu, and Z. W. Xing. 2005. "The Effect of China's One-Child Family Policy after 25 Years." *The New England Journal of Medicine* 353: 1171–1176. doi:10.1056/NEJMhpr051833.

Johnson, E., and G. Chow. 1997. "Rates of Return to Schooling in China." *Pacific Economic Review* 2: 101–113. doi:10.1111/per.1997.2.issue-2.

Lall, A., and C. Sakellariou. 2010. "Evolution of Education Premiums in Cambodia: 1997–2007." *Asian Economic Journal* 24 (4): 333–354. doi:10.1111/j.1467-8381.2010.02044.x.

Li, H. 2003. "Economic Transition and Returns to Education in China." *Economics of Education Review* 22 (3): 317–328. doi:10.1016/S0272-7757(02)00056-0.

Li, H., and Y. Luo. 2004. "Reporting Errors, Ability Heterogeneity and Returns to Schooling in China." *Pacific Economic Review* 9 (3): 191–207. doi:10.1111/per.2004.9.issue-3.

Liu, Z. 1998. "Earnings, Education and Economic Reforms in Urban China." *Economic Development and Cultural Change* 46 (4): 697–725. doi:10.1086/edcc.1998.46.issue-4.

Lu, Z., and S. Song. 2006. "Rural-Urban Migration and Wage Determination: The Case of Tianjin, China." *China Economic Review* 17: 337–345. doi:10.1016/j.chieco.2006.04.007.

Mayer, S., and D. Knutson. 1999. "Does the Timing of School Affect How Much Children Learn?" In *Earning and Learning: How Schools Matter*, ed. Mayer, S. and P. Peterson, 79–102. Washington, DC: Brookings Institution Press.

Meng, X., and M. P. Kidd. 1997. "Labor Market Reform and the Changing Structure of Wage Determination in China's State Sector during the 1980s." *Journal of Comparative Economics* 25 (3): 403–421. doi:10.1006/jcec.1997.1481.

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

Murray, M. P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4): 111–132. doi:10.1257/jep.20.4.111.

Pencavel, J. 1998. "Assortative Mating by Schooling and the Work Behaviour of Wives and Husbands." *American Economic Review* 88 (2): 326–329.

Puhani, P., and A. Weber. 2005. "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany." *IZA Discussion Paper No. 1827*. Bonn, Germany: Institute for the Study of Labour.

Qian, N. 2009. "Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrolment in Rural China." *Working Paper 14973*, Cambridge, MA: National Bureau of Economic Research.

Rodgers, J. L., H. H. Cleveland, E. Van Den Oord, and D. C. Rowe. 2000. "Resolving the Debate Over Birth Order, Family Size and Intelligence." *American Psychologist* 55 (6): 599–612. doi:10.1037/0003-066X.55.6.599.

Stock, J., and M. Yogo. 2005. "Testing for Weak Instruments in IV Regression." In *Identification and Inference for Econometric Models: A Festschrift in Honour of Thomas Rothenberg*, eds. Donald, W. K. A. and J. H. Stock, 80–108. Cambridge, MA: Cambridge University Press.

Strøm, B. 2004. *Student Achievements and Birthday Effects*. Trodheim, Norway: Mimeo, Norwegian University of Science and Technology.

Trostel, P., I. Walker, and P. Woolley. 2002. "Estimates of the Economic Return to Schooling for 28 Countries." *Labour Economics* 9 (1): 1–16. doi:10.1016/S0927-5371(01)00052-5.

Whichman, A. L., J. L. Rodgers, and R. C. McCallum. 2006. "A Multi-Level Approach to the Relationship between Birth Order and Intelligence." *Personality and Social Psychology Bulletin* 32 (1): 117–127. doi:10.1177/0146167205279581.

World Development Indicators. 2009. *The World Bank*. Washington, DC: The World Bank.

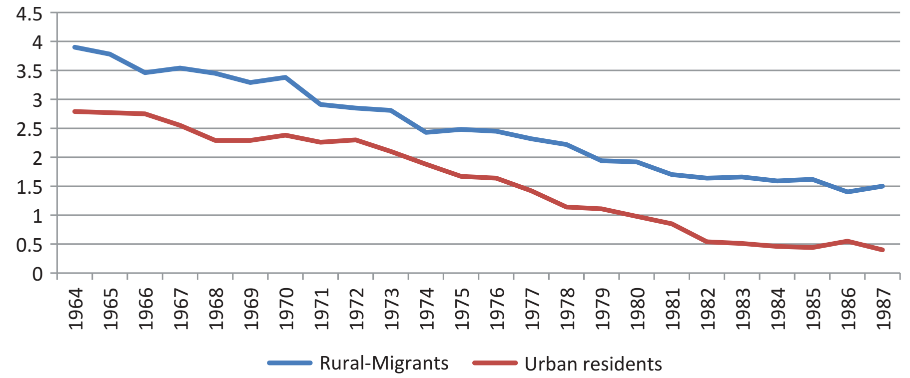
Zajonc, R. B. 1976. "Family Configuration and Intelligence." *Science* 192: 227–236. doi:10.1126/science.192.4236.227.

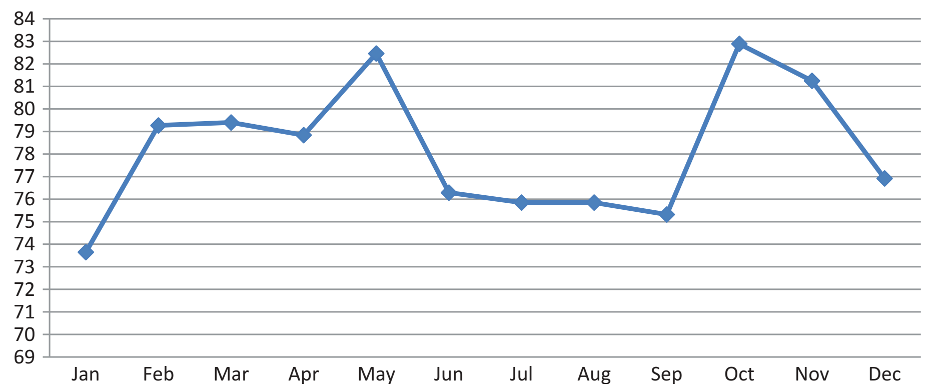
Zajonc, R. B., and G. B. Markus. 1975. "Birth Order and Intellectual Development." *Psychological Review* 82: 74–88. doi:10.1037/h0076229.

Zajonc, R. B., and P. R. Mullally. 1997. "Birth Order: Reconciling Conflicting Effects." *American Psychologist* 52: 685–699. doi:10.1037/0003-066X.52.7.685.

Zhang, J., Y. Zhao, A. Park, and X. Song. 2005. "Economic Returns to Schooling in Urban China, 1988 to 2001." *Journal of Comparative Economics* 33: 730–752. doi:10.1016/j.jce.2005.05.008.

Appendix


4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987
Rural-Migrants
Urban residents


84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

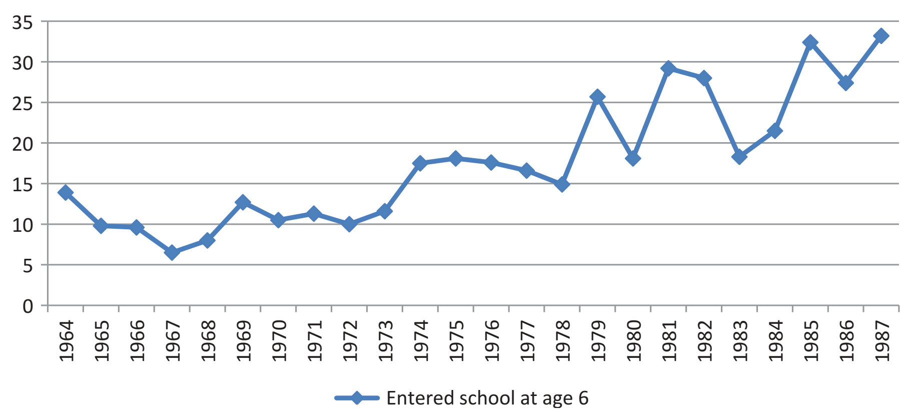

35
30
25
20
15
10
5
0
1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987
Entered school at age 6

Table A1. Married male urban workers.

	OLS[a]	IV	IV	IV
Years of schooling	**0.072**	**0.095**	**0.094**	**0.134**
	(0.007)	(0.013)	(0.012)	(0.034)
Tenure	**0.029**	**0.033**	**0.033**	**0.043**
	(0.008)	(0.008)	(0.008)	(0.010)
Tenure squared	**−0.0007**	**−0.0007**	**−0.0007**	**−0.0008**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Other experience	0.009	**0.017**	**0.017**	**0.031**
	(0.007)	(0.007)	(0.007)	(0.013)
Other experience squared	**−0.0006**	**−0.0007**	**−0.0007**	**−0.0008**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Firm size: 6–20	**0.183**	**0.160**	**0.161**	0.122
	(0.076)	(0.068)	(0.068)	(0.077)
Firm size: 21–100	**0.275**	**0.245**	**0.246**	**0.210**
	(0.069)	(0.063)	(0.063)	(0.078)
Firm size: >100	**0.265**	**0.236**	**0.237**	**0.190**
	(0.067)	(0.060)	(0.060)	(0.074)
Firstborn	0.002	–	–	–
	(0.035)			
Constant	**1.39**	**1.01**	**1.03**	0.322
	(0.154)	(0.232)	(0.222)	(0.557)
First stage				
Shea partial R^2/partial R^2		0.272	0.299	0.045
F-Value [*p*-value]		641.0 [0.000]	244.2 [0.000]	41.74 [0.000]
Second stage				
Over-identification test: all instruments:				
Sargan statistic [*p*-value]		–	0.090 [0.956]	0.184 [0.668]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for				
- Spouse's years of schooling		–	0.089 [0.766]	–
- Firstborn		–	0.089 [0.766]	–
- Firstborn interacted with age		–	0.090 [0.764]	–
Weak identification test				
Cragg–Donald Wald statistic (*F*-statistic)		641.0	244.2	0.045
Stock–Yogo weak ID test critical values				
5% maximal IV relative bias		–	13.91	
10% maximal IV size		16.38	22.30	19.93
Redundancy test (χ^2 [p-value] for				
- Spouse's years of schooling		–	461.4 [0.000]	–
- Firstborn		–	64.00 [0.000]	75.87 [0.000]
- Firstborn interacted with age		–	66.18 [0.000]	79.96 [0.000]
Endogeneity test for years of schooling				
χ^2 [*p*-value]		4.39 [0.036]	4.54 [0.033]	2.98 [0.084]
R^2	0.210			
N	1725	1725	1725	1725

Note: [a]Robust SEs in parentheses; OLS results using the same sample as for IV estimation. Significant at the 5% level or lower in bold.

Table A2. Married female urban workers.

	OLS[a]	IV	IV	IV	IV
Years of schooling	**0.057**	**0.090**	**0.080**	**0.066**	**0.068**
	(0.008)	(0.014)	(0.011)	(0.015)	(0.014)
Tenure	0.009	0.015	0.013	0.011	0.011
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Tenure squared	−0.0003	−0.0003	−0.0003	−0.0003	−0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Other experience	−0.012	0.002	−0.002	−0.008	−0.007
	(0.008)	(0.009)	(0.008)	(0.009)	(0.009)
Other experience squared	−0.0001	−0.0003	−0.0002	−0.0002	−0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Firm size: 6–20	**0.135**	0.082	0.095	**0.124**	**0.120**
	(0.063)	(0.063)	(0.062)	(0.063)	(0.063)
Firm size: 21–100	**0.344**	**0.303**	**0.314**	**0.337**	**0.331**
	(0.064)	(0.061)	(0.060)	(0.061)	(0.061)
Firm size: >100	**0.398**	**0.361**	**0.373**	**0.390**	**0.385**
	(0.061)	(0.059)	(0.058)	(0.059)	(0.059)
Has brother/s	−0.048	–	-0.042	−0.043	−0.047
	(0.049)		(0.047)	(0.047)	(0.047)
Firstborn	−0.039	–	–	–	−0.043
	(0.037)				(0.040)
Constant	**1.45**	**0.928**	**1.08**	**1.30**	**1.30**
	(0.162)	(0.235)	(0.201)	(0.252)	(0.243)
First stage					
Shea partial R^2/partial R^2		0.309	0.450	0.260	0.278
F-Value [*p*-value]		569.0 [0.000]	346.4 [0.000]	224.7 [0.000]	143.6 [0.000]
Second stage					
Over-identification test: all instruments					
Sargan statistic [*p*-value]		–	1.78 [0.410]	0.140 [0.708]	0.758 [0.685]
Exogeneity/orthogonality of suspect instruments (C-test [p-value]) for					
- Spouse's years of schooling		–	1.54 [0.215]		–
- Has brother/s interacted with age		–	1.56 [0.212]	–	0.127 [0.722]
- Has brother/s interacted with number of siblings		–	0.079 [0.778]	–	0.529 [0.467]
- Firstborn interacted with age		–	–	–	0.175 [0.676]
Weak identification test		569.0	346.4		143.6
Cragg–Donald Wald statistic (*F*-statistic)				224.7	
Stock–Yogo weak ID test critical values		–	13.91		13.91
5% maximal IV relative bias		16.38	22.30	–	22.30
10% maximal IV size		–		19.93	
Redundancy test (χ^2 [p-value] for					
- Spouse's years of schooling		–	332.4 [0.000]	–	–
- Has brother/s interacted with age		–	261.0 [0.000]	335.1 [0.000]	314.9 [0.000]
- Has brother/s interacted number of siblings		–	3.73 [0.054]	10.60 [0.001]	4.43 [0.035]
- Firstborn interacted with age		–	–	–	32.33 [0.000]
Endogeneity test for years of schooling		7.94 [0.005]	7.00 [0.008]	0.545 [0.460]	0.796 [0.372]
χ^2 [*p*-value]					
R^2	0.247				
N	1282	1282	1282	1282	1282

Note: [a]Robust SEs in parentheses; OLS results using the same sample as for IV estimation. Significant at the 5% level or lower in bold.