

# Returns to schooling for urban residents and migrants in China: New IV estimates and a comprehensive investigation

Chris Sakellariou  
Division of Economics  
Humanities and Social Sciences  
Nanyang Technological University  
Singapore  
E-mail: [acsake@ntu.edu.sg](mailto:acsake@ntu.edu.sg)

Zheng Fang  
Southwestern University of Finance and Economics  
China  
E-mail: [fangzheng@swufe.edu.cn](mailto:fangzheng@swufe.edu.cn)

**Abstract:** This paper uses a new dataset, the 2009 Rural Urban Migration in China (RUMiC) to estimate returns to schooling in China using an instrumental variable (IV) methodology. 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 estimates are in a fairly tight band for all four sub-samples examined (urban men, urban women, migrant men and migrant women). Estimates for men range from about 9.5% for urban workers to about 10-10.5% for migrant workers and are slightly higher than the corresponding estimates for women, which range from 7.5% for female urban workers to 8-9.5% for female migrant workers. Thus, private returns to education in urban China in 2009 were substantial and of similar magnitude to those for other transition countries, as well as to worldwide and developing country averages. 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.

*JEL Classification: I21*

*Keywords: Returns to Schooling, Instrumental Variables, Rural-to-Urban Migrants, China*

## 1. 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 3-5% 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 one to deal 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 that more educated workers migrate? Are returns for women higher than those of men (as most past studies seem to indicate), or comparable to those of men?

This paper uses the most recent data available, the 2009 Rural Urban Migration in China (RUMiC), a rich dataset which allows exploration for potentially suitable instruments to 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. We find that estimated returns are in a fairly tight band for all four sub-samples examined (urban men, urban women, migrant men and migrant women). Estimates for men range from about 9.5% for urban workers to about 10-10.5% for migrant workers; corresponding estimates for women range from 7.5% for female urban workers to 8-9.5% for female migrant workers. We, thus, find that returns for men are slightly higher than returns for women and migrant returns are slightly higher than returns for urban workers.

## **2. Literature Review**

Past research on returns to education in China varies in focus, examining issues such as the effect of economic reforms on returns, differences in urban, rural and migrant returns, and methodological issues. 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 world average or the average for

the region (both of which are about 10%). We follow with a non-exhaustive 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 *et. al.* (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 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 (see for example, Li 2003). A more recent study by Zhang *et. al.* (2005) used the Mincerian approach and focused on returns to schooling in urban China over an extended period of economic reforms and rising income inequality. They find a dramatic increase in the returns to education, from only 4.0 percent per year of schooling in 1988 to 10.2 percent 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 different methodological approach. They find that returns are higher than those reported earlier at about 6.5 percent and even higher for younger people and migrants.

Some more recent studies try to address the omitted ability and measurement error biases using mostly instrumental variable (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 instrumental variable to address ability heterogeneity. They find returns to schooling much higher than those from OLS, at about 15%, for young workers in China. Chen and Hamori (2008) 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% respectively for married men and women. The estimate for the return from instrumental variable 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 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 do not attend. They conclude that the return to education has increased substantially in China since the early 1990s.

### **3. Data**

The Survey on Rural Urban Migration in China (RUMiC) was established 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. There is particular emphasis on the welfare status of migrants, i.e., 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<sup>1</sup>, and the Urban Migrant Survey covers 15 cities (which are provincial capital cities or other major migrant receiving cities)<sup>2</sup> in nine provinces or metropolitan areas. The Urban Household Survey was conducted in 19 cities.<sup>3</sup> 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.

## 4. Methodology

### 4.1 Looking for an instrument set

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

$$\text{Ln}(W_i) = \beta_0 + \beta_1 Y_i + \beta_2 X_i + \varepsilon_i$$

where  $W$  is the wage rate,  $Y$  is years of education completed,  $X$  is a vector of other controls assumed to be exogenous and  $\varepsilon$  is the error term. Instrumental variables estimation entails identifying a set of variables  $Z$  (the set of instruments) which: is uncorrelated with  $\varepsilon$ ; 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:

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

are used to derive the fitted values of  $Y_i$  using ordinary least squares (OLS).

The OLS 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

---

<sup>1</sup>These are: Anhui, Chongqing, Guangdong, Hebei, Henan, Hubei, Jiangsu, Sichuan, and Zhejiang.

<sup>2</sup>These are: Bengbu, Chengdu, Chongqing, Dongguan, Guangzhou, Hefei, Hangzhou, Luoyang, Nanjing, Ningbo, Shanghai, Shenzhen, Wuhan, Wuxi, Zhengzhou.

<sup>3</sup> It includes the following additional cities to the Urban Migrant Survey: Anyang, Jiande, Leshan and Mianyang.

with the disturbance term. The instrumental variable estimator (by being a consistent estimator) can avoid the bias that ordinary least squares suffers from, when an explanatory variable in a regression is correlated with the regression's disturbance term; however, instrumental variable 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 isn't "too weak" (i.e., it is sufficiently correlated with the endogenous explanatory variable). In what follows in this sub-section, we outline the main considerations guiding our methodological approach<sup>4</sup> in estimating the return to schooling for rural to urban migrants and urban residents in China using instrumental variable 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 an IV estimator, such as the two-stage least squares estimator, the standard errors are larger 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 (such as Sargan's test) 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

---

<sup>4</sup> Murray (2006) provides an excellent discussion for what one should consider when assessing validity and strength of instruments.

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 instruments 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, standard errors 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<sup>5</sup>. 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

---

<sup>5</sup>In the Stock-Yogo test, the hypothesis is that the IV bias is less than some fraction (say 5%) of the OLS bias.

the head of household, if the data does not contain information on education of parents for all individuals).

#### 4.2 *The instrument set*

After extensive exploration, the instrument set used is based on the following information available in the RUMiC 2009 surveys: 1. Birth order information and in particular being firstborn vs. later-born, along with the interaction with age cohort and/or number of siblings, depending on estimation sample (migrants by gender and urban residents by gender); 2. Age of school entry: age 6 vs. age 7 (available only in the Migrant survey), along with its interaction with age cohort; 3. Early smoking history, which entered the instrument set only in male regressions; 4. Spouse's years of schooling completed, which is used only for robustness checks of our estimates. Below we discuss each of the instruments used:

**Firstborn:** There is an emerging literature exploring the degree to which family size 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 we control for birth order. Fergusson *et.al.* (2006) used New Zealand data and nested models to control for the confounding effects of family size on birth order and found that 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 find that birth order has a strong negative effect on a child's education, consistent with existing empirical studies. Overall, they provide 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 finds 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 child 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. Families in which neither parent has siblings are also allowed to have two

children.<sup>6</sup> 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; some provinces had other exemptions worked into their policies as well (Sichuan, for example, has 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 *et. al.*, 2005). Chart 1 in the appendix depicts over-time declines in the average number of siblings for adults in the 22-45 age group (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 a 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 rather small. Several such papers

---

<sup>6</sup> 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 last century, 27 provinces and municipalities have passed it; but it achieved the full coverage in China only until 2011 when Henan government finally agreed to the implementation of such a policy. In order to address the demographic challenges such as aging population and shrinking labour force, the one-child policy was relaxed further in the Third Plenary Session of the 18<sup>th</sup> CPC Central Committee in 2013, where families with one parent being the only child would soon be permitted to have two children (the time of implementation is subject to the revision of the regulation of the local government).

subsequently appeared leading to the “confluence” hypothesis (for example Zajonc and Markus, 1975; Zajonc, 1976; Zajonc & 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 siblings were compared to one another. Whichman *et. al.* (2006) observe 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 (NLSY) data and compared siblings to one-another at fixed ages and found conclusive evidence that the fundamental cause of presumed birth order effects 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.

**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 dataset 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 (that is, 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 attain a low schooling degree or less in the Eastern part of the country, i.e., 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 consider it is likely that these findings could be driven by unobserved heterogeneity, i.e. those individuals who entered late did so since 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 seven is associated with a better school performance. Puhani and Weber (2005) used three German datasets and instrumental variables estimation and found robust and significant positive effects on educational attainment for pupils who enter school at seven instead of six years of age; in particular, they find that test scores at the end of primary school increase by about 0.42 standard deviations 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 and our estimates 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 September 1st and sixteenth of the first month of Chinese lunar year, with a summer vacation in July and August and a winter vacation around Chinese spring festival; however, over the years there were periods with a spring enrolment (April). Our estimation sample includes those in the 22-45 age group (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. 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.

According to the provisions of the Compulsory Education Law of the People's Republic of China, the six years of primary education start at age six (fully six years old) or seven; children usually entered primary school at seven 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 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.

**Early smoker:** The rationale for using early smoking behaviour as an instrument is provided by Evans and Montgomery (1994) who argued that the correlation between education and health was due to the unobserved differences in the discount rates or time preference across individuals. Investments in both education and health involve a trade-off between current costs and future benefits. Smoking is indicative of the fact that individuals who smoke show

that they place considerable weight on satisfying current wants at the expense of future benefits. Early smoking behaviour is not correlated with current earnings but is correlated with educational choices because these are also made in the mid to late teens. The scarce prior empirical evidence that uses smoking as an instrument to estimate returns to schooling (Fersterer and Winter-Ebmer (2003); Harmon et al. 2000; Lall and Sakellariou 2010) finds larger estimated returns to education from IV estimation than from OLS estimation. In this paper, the binary instrument takes the value of 1 for (male) smokers who started smoking before the age of 18, and 0 for non-smokers and those who started smoking later.

**Spouse's education:** This instrument was used to evaluate the robustness of the main findings. Spouse's education is suggested as a possible valid instrument by Trostel et al. (2002) who explore 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 et al. (2002) obtained estimates using spouse's education to instrument for schooling that are over 20 percent 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 find higher returns to schooling from two stage least squares compared to OLS, especially for women.

## **5. Estimation and Results**

### *5.1 Results*

#### Migrant workers

The estimation sample used is for workers 22-45 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 includes 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 were 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 women only) and instrumental variable 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). Information on smoking history is available for 99% of persons in the migrant sample and all persons in the urban sample.

Table 1 presents the results for male migrant workers. Column 1 gives the OLS estimates for the entire sample, column 2 the OLS estimates using the same sample as for the IV estimation and columns 3 and 4 the IV estimates for different combinations of instruments. OLS estimates of the return to schooling are in the order of only 4-5%. The full set of instruments available includes the interaction terms of firstborn with age and with number of siblings; it also includes the interaction of age at school entry with age (given the increasing proportion of schoolchildren entering school at age 6 rather than age 7 over time).

In column 3, estimates are based on a full set of available instruments (with the exception of firstborn interacted with number of siblings, which was redundant). The estimate of the return to schooling, at just over 10% is more than twice the OLS estimate and

statistically significant at the 1% level. Return to an additional year of tenure is estimated at just over 3%, while the estimate of the return to an additional year of other experience is small and imprecisely estimated. The instrument set is strong with a Shea/partial  $R^2$  of 0.082 and F-value of just over 20. This value of F-statistic exceeds the critical value for 5% relative bias and 15% maximal IV size (suggesting that the inference on estimated standard errors is most likely valid). All included instruments pass the redundancy test; Firstborn and its interaction with age are the relatively stronger instruments. Concerning validity of the instrument set, the Sargan statistic test shows that the instrument set is valid (p-value nearly 1); likewise, testing exogeneity of individual instruments shows that all C-statistics are associated with very high p-values. Finally, the endogeneity test for years of schooling suggests that it is likely endogenous.

In column 4 we evaluate the robustness of estimates by excluding the instruments based on age of school entry. The estimate of the return to schooling does not change much - at just under 11%. Similarly, the estimates for the return to tenure and other controls remain approximately unchanged. The instrument set remains strong with partial  $R^2$  at 0.057, while the F-value is higher, at 28.4. The value of the F-statistic exceeds the critical value for 5% relative bias and 10% maximal IV size. The Sargan statistic and the C-statistics confirm the validity of the instrument set as well as each individual instrument.

Table 1: Male migrant workers, age 22-45

	OLS	OLS <sup>1</sup>	IV <sup>2</sup>	IV <sup>3</sup>
Years of schooling	0.042*** (0.008)	0.045*** (0.009)	0.101*** (0.032)	0.107*** (0.036)
Tenure	0.039*** (0.009)	0.035*** (0.009)	0.042*** (0.012)	0.042*** (0.011)
Tenure squared	-0.0014*** (0.0005)	-0.0013*** (0.0005)	-0.0011* (0.0007)	-0.0012** (0.0006)
Other experience	-0.015* (0.008)	-0.015* (0.009)	0.0023 (0.012)	0.0022 (0.013)
Other experience squared	0.0003 (0.0003)	0.0003 (0.0003)	0.0000 (0.0003)	0.0001 (0.0003)
Married	0.093*** (0.034)	0.101*** (0.037)	0.055 (0.055)	0.054 (0.053)
Firm size: 6-20	0.188*** (0.049)	0.209*** (0.055)	0.190*** (0.057)	0.211*** (0.054)
Firm size: 21-100	0.286*** (0.047)	0.303*** (0.053)	0.307*** (0.057)	0.292*** (0.053)
Firm size: > 100	0.382*** (0.045)	0.384*** (0.050)	0.373*** (0.050)	0.356*** (0.047)
Constant	1.27*** (0.112)	1.25*** (0.125)	0.566 (0.387)	0.513 (0.432)
<u>First stage:</u>				
Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>			0.082	0.057
F-value [p-value]			20.14 [0.000]	28.43 [0.000]
<u>Second stage</u>				
<i>Over-identification test: all instruments:</i>				
Sargan statistic [p-value]			0.130 [0.998]	0.330 [0.848]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>				
-	Early smoker		0.001 [0.980]	0.058 [0.810]
-	Firstborn		0.048 [0.827]	0.063 [0.803]
-	Firstborn interacted with age		0.033 [0.855]	0.006 [0.940]
-	Started school at age 6 (vs. at age 7)		0.003 [0.957]	-
-	Started school at age 6 interacted with age		0.000 [0.992]	-
<i>Weak identification test</i>				
Cragg-Donald Wald statistic (F-statistic):			20.14	28.43
Stock-Yogo weak ID test critical values:				
5% maximal IV relative bias			18.37	13.91
10% maximal IV size			26.87	22.30
15% maximal IV size			15.09	-
<i>Redundancy test (Chi-sq [P-value]) for:</i>				
-	Early smoker		12.34 [0.000]	15.71 [0.000]
-	Firstborn		52.21 [0.000]	51.77 [0.000]
-	Firstborn interacted with age		61.19 [0.000]	61.20 [0.000]
-	Started school at age 6 (vs. at age 7)		10.68 [0.001]	-
-	Started school at age 6 interacted with age		14.09 [0.000]	-
<i>Endogeneity test for years of schooling</i>				
Chi-sq [P-value]			2.74 [0.098]	3.24 [0.072]
R <sup>2</sup>	0.121	0.119		
N	1,678	1,416	1,145	1,416

Note: robust standard errors in parentheses. <sup>1</sup>OLS results using the same sample as for IV estimation in column 4. <sup>2</sup>Instruments: Started school before age 7, started school before age 7 interacted with age, early smoker, firstborn, firstborn interacted with age. <sup>3</sup>Instruments: early smoker, firstborn, firstborn interacted with age.

Table 2 presents the results for female migrants. Columns 1 and 2 contain the OLS estimates, while column 3 contains selectivity corrected estimates using Heckman's correction procedure. The estimate of the return to schooling from OLS is less than 4% (at 3-3.5%) and the corresponding selectivity corrected estimate is not much different at 3% (with the inverse Mills ratio being insignificant). Column 4 contains the main IV estimates. Compared to estimation for men, here the instrument "early smoker" is excluded and the interaction of firstborn with number of siblings enters the instrument set. The estimate of the return to schooling at 9.5% is three times the OLS estimate and statistically significant at the 1% level. The estimate of the return to tenure, at about 4.5%, is higher than the corresponding estimate for men. The return to other experience, at about 2% is significant at the 10% level.

As was the case for estimates for men, the instrument set is not weak, with a partial  $R^2$  of 0.10; the F-statistic value of 17.8 exceeds the critical value for 10% relative bias and 15% maximal IV size, while all instruments are relevant. Based on the value of the Sargan statistic, the instrument set is valid; the C-statistics do not reject the hypothesis that each individual instrument can be considered exogenous.

When the instruments based on age of school entry are omitted (column 4), the estimate of the return to schooling decreases by about 1 percentage point to just over 8% (statistically significant at the 5% level). Checking the instrument strength using the Stock-Yogo statistics, the F-statistic value of 19.1 exceeds the critical value for 5% relative bias and 15% maximal IV size. The Sargan statistic p-value is similar to that in the main estimates and the C-statistic values confirm exogeneity of individual instruments.

Table 2: Female migrant workers, age 22-45

	OLS	OLS <sup>1</sup>	Heckman <sup>2</sup>	IV <sup>3</sup>	IV <sup>4</sup>
Years of schooling	0.035*** (0.007)	0.033*** (0.008)	0.029*** (0.008)	0.091*** (0.030)	0.081** (0.036)
Tenure	0.054*** (0.011)	0.055*** (0.011)	0.051*** (0.011)	0.060*** (0.013)	0.064*** (0.013)
Tenure squared	-0.0024*** (0.0007)	-0.0024*** (0.0007)	-0.0022*** (0.0007)	-0.0021*** (0.0008)	-0.0025*** (0.0007)
Other experience	-0.0005 (0.008)	-0.0018 (0.008)	-0.0029 (0.009)	0.021* (0.013)	0.010 (0.012)
Other experience squared	-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0002 (0.0003)	-0.0006* (0.0003)	-0.0002 (0.0003)
Married	-0.100** (0.040)	-0.086** (0.041)	-0.052 (0.050)	-0.041 (0.055)	-0.072 (0.049)
Firm size: 6-20	0.114** (0.045)	0.111** (0.048)	0.105** (0.050)	0.083 (0.060)	0.084 (0.054)
Firm size: 21-100	0.247*** (0.048)	0.289*** (0.050)	0.291*** (0.050)	0.292*** (0.059)	0.262*** (0.054)
Firm size: > 100	0.230*** (0.041)	0.251*** (0.042)	0.250*** (0.044)	0.227*** (0.053)	0.225*** (0.048)
Constant	1.29*** (0.106)	1.27*** (0.111)	1.42*** (0.143)	0.450 (0.376)	0.693 (0.437)
lambda			-0.518 (0.328)		
<u>First stage:</u> Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>				0.100	0.052
F-value [p-value]				17.81 [0.000]	19.06 [0.000]
<u>Second stage</u>					
<i>Over-identification test: all instruments:</i>					
Sargan statistic [p-value]				2.47 [0.650]	1.02 [0.601]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>					
- Entered school at age 6 (vs age 7)				0.554 [0.457]	-
- Entered school at age 6 interacted with age				0.287 [0.592]	-
- Firstborn				0.001 [0.980]	0.176 [0.675]
- Firstborn interacted with age				0.011 [0.917]	0.103 [0.748]
- Firstborn interacted with number of siblings				0.166 [0.684]	0.928 [0.336]
<i>Weak identification test</i>					
Cragg-Donald Wald statistic (F-statistic):				17.81	19.06
<i>Stock-Yogo weak ID test critical values:</i>					
5% maximal IV relative bias				18.37	13.91
10% maximal IV relative bias				26.87	22.30
15% maximal IV size				15.09	12.83
<i>Redundancy test (Chi=sq. [P=value]) for:</i>					
- Entered school at age 6 (vs age 7)				11.76 [0.001]	-
- Entered school at 6 interacted with age				18.96 [0.000]	-
- Firstborn				22.29 [0.000]	25.41 [0.000]
- Firstborn interacted with age				37.69 [0.000]	44.76 [0.000]
- Firstborn interacted with number of siblings				12.13 [0.000]	14.07 [0.000]
<i>Endogeneity test for years of schooling</i>					
Chi-sq [P-value]:				3.22 [0.073]	1.98 [0.160]
R <sup>2</sup>	0.122	0.125			
Wald chi2 [P-value]			124.3 [0.000]		
N	1,150	1,062	1,168 (censored 137)		1,062

Note: robust standard errors in parentheses. <sup>1</sup>OLS results using the same sample as for IV estimation. <sup>2</sup>Independent variables in selection equation: age, age squared, marital status, number of children, majority ethnic group (same sample as IV estimation). <sup>3</sup>Instruments: entered school at age 6, entered school at age 6 interacted with age, firstborn, firstborn interacted with age. <sup>4</sup>Instruments: firstborn, firstborn interacted with age and firstborn interacted with number of siblings.

### Urban workers

The instrument set for estimating returns to schooling for urban workers is based on sibling composition only (with the addition of early smoking behaviour for men only), since information on age of school entry is not available in the urban survey. Table 3 presents the results for men. OLS estimates of the coefficient of years of schooling are higher compared to this for migrants, at about 6%. The estimate of the return to schooling from IV estimation is in the same order of magnitude as for migrant workers, at approximately 10%. The OLS-IV gap in estimates is much smaller for the urban sample; this is likely due to a smaller measurement error bias in the urban sample. The instrument set is strong, with an F-value of about 50, 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. The endogeneity test for years of schooling indicates that it is endogenous; the p-value is generally smaller compared to that for male migrants despite the smaller OLS-IV estimate gap, as the sample size is larger and estimates more precise.

The return to an additional year of tenure is concave, with earnings increasing by about 2.5% per additional year. The premium associated with being married is significantly higher than the premium for male migrant workers and exceeds 25%. Finally, as was the case with the results for migrant workers, there is a significant premium associated with working in larger firms.

The results for women (see Table 4) are similar to those for men; however, OLS and IV estimates of the return to schooling are even closer than in Table 3; the IV estimate at 7.5% is just 1.5 percentage points higher than the OLS and selectivity corrected estimates (at about 6%). Once again, the instrument set is strong and there are no major issues with the validity of the instrument set (although the C-statistic p-value is somewhat low for one of the

instruments in the set). In this case, 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 in the appendix). Column 1 gives the OLS estimates for the same sample as the derived IV estimates. Columns 2-4 give the IV estimates with spouse's education as the sole instrument (column 2), a combination of spouse's education, firstborn, firstborn interacted with age and history of early smoking (column 3), and the same combination of instruments with the exclusion of spouse's education (column 4). Estimates for married male urban workers are remarkably close to those in the main results. The three IV estimates of the return to schooling are within less than 1 percentage point of one another and also within less than 1 percentage point of the main IV estimate in Table 3. Similar robustness checks using married female urban workers, show that the IV estimates in columns 3 and 4, at 7.5%, are identical to the estimate in the main results (in Table 4), while the estimate in column 2 using only spouse's education as instrument is about 1 percentage point higher at about 8.5%. Instrument sets are strong and tests validate the instruments in column 3, while in column 4, the Sargan statistic p-value, at about 0.15, is somewhat low.

Table 3: Male urban workers, age 22-45

	OLS	OLS <sup>1</sup>	IV <sup>2</sup>
Years of schooling	0.066*** (0.007)	0.054*** (0.007)	0.095*** (0.023)
Tenure	0.050*** (0.008)	0.044*** (0.010)	0.053*** (0.011)
Tenure squared	-0.0015*** (0.0003)	-0.0014*** (0.0003)	-0.0014*** (0.0003)
Other experience	0.010 (0.007)	0.014* (0.008)	0.033** (0.014)
Other experience squared	-0.0007** (0.0003)	-0.0010*** (0.0003)	-0.0013*** (0.0004)
Married	0.221*** (0.050)	0.325*** (0.062)	0.254*** (0.072)
Firm size: 6-20	0.214*** (0.069)	0.237*** (0.076)	0.177** (0.075)
Firm size: 21-100	0.294*** (0.065)	0.337*** (0.071)	0.269*** (0.072)
Firm size: > 100	0.265*** (0.063)	0.285*** (0.069)	0.223*** (0.067)
Constant	1.06*** (0.119)	1.13*** (0.138)	0.517 (0.358)
<u>First stage:</u>			
Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>			0.081
F-value [p-value]			49.61 [0.000]
<u>Second stage</u>			
<i>Over-identification test: all instruments:</i>			
Sargan statistic [p-value]			1.14 [0.566]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>			
- Early smoker			0.006 [0.936]
- Firstborn			0.538 [0.463]
- Firstborn interacted with age			0.183 [0.669]
<i>Weak identification test</i>			
Cragg-Donald Wald statistic (F-statistic):			49.61
Stock-Yogo weak ID test critical values:			
5% maximal IV relative bias			13.91
10% maximal IV size			22.30
<i>Redundancy test (Chi=sq. [P=value]) for:</i>			
- Early smoker			12.27 [0.000]
- Firstborn			112.2 [0.000]
- Firstborn interacted with age			121.1 [0.000]
<i>Endogeneity test for years of schooling</i>			
Chi-sq. [P-value]			3.49 [0.062]
R <sup>2</sup>	0.169	0.177	
N	2,233	1,703	1,703

Note: robust standard errors in parentheses. <sup>1</sup>OLS results using the same sample as for IV estimation. <sup>2</sup>Instruments: Early smoker, firstborn, firstborn interacted with age.

Table 4: Female urban workers, age 22-45

	OLS	OLS <sup>1</sup>	Heckman <sup>2</sup>	IV <sup>3</sup>
Years of schooling	0.056*** (0.007)	0.052*** (0.008)	0.060*** (0.007)	0.075*** (0.022)
Tenure	0.038*** (0.007)	0.032*** (0.007)	0.034*** (0.009)	0.037*** (0.010)
Tenure squared	-0.0013*** (0.0003)	-0.0012*** (0.0003)	-0.0012*** (0.0003)	-0.0012*** (0.0003)
Other experience	-0.014* (0.007)	-0.0017** (0.008)	-0.014* (0.009)	-0.006 (0.013)
Other experience squared	0.0002 (0.0003)	0.0002 (0.0003)	0.0002 (0.0003)	0.0000 (0.0003)
Married	0.084* (0.043)	0.114* (0.060)	0.093 (0.087)	0.100* (0.058)
Firm size: 6-20	0.026 (0.060)	0.064 (0.062)	0.067 (0.060)	0.028 (0.067)
Firm size: 21-100	0.228*** (0.059)	0.287*** (0.061)	0.293*** (0.056)	0.259*** (0.060)
Firm size: > 100	0.242*** (0.057)	0.282*** (0.058)	0.285*** (0.054)	0.250*** (0.060)
Constant	1.27*** (0.119)	1.33*** (0.136)	0.936*** (0.285)	0.962*** (0.357)
lambda			0.711 (0.512)	
<u>First stage:</u>				
Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>				0.085
F-value [p-value]				52.67 [0.000]
<u>Second stage</u>				
<i>Over-identification test: all instruments:</i>				
Sargan statistic [p-value]				2.08 [0.353]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>				
- Firstborn				0.552 [0.458]
- Firstborn interacted with age				1.73 [0.188]
- Firstborn interacted with # of siblings				0.397 [0.529]
<i>Weak identification test</i>				
Cragg-Donald Wald statistic (F-statistic):				52.67
Stock-Yogo weak ID test critical values:				
10% maximal IV size				22.30
<i>Redundancy test (Chi=sq. [P=value]) for:</i>				
- Firstborn				126.1 [0.000]
- Firstborn interacted with age				142.7 [0.000]
- Firstborn interacted with # of siblings				5.78 [0.016]
<i>Endogeneity test for years of schooling</i>				
Chi-sq. [P-value]				1.25 [0.263]
R <sup>2</sup>	0.186	0.193		
Wald chi <sup>2</sup> [P-value]			412.1 [0.000]	
N	2,184	1,719	2,048 (censored 439)	1,719

Note: robust standard errors in parentheses.<sup>1</sup>OLS results using the same sample as for IV estimation. <sup>2</sup>Independent variables in selection equation: age, age squared, marital status, number of children, majority. <sup>3</sup>Instruments: firstborn, firstborn interacted with age, firstborn interacted with number of siblings.

Summarizing, this paper examines the economic returns to education in urban China (including returns for migrant workers) with particular interest in instrumental variable (IV) estimation; we paid particular attention to the evaluation of instrument sets used with respect to validity and relevance. Using 2009 Survey on Rural Urban Migration in China data, we identify and use a set of instruments in various combinations. This allows us to go beyond merely providing first-stage statistics and testing over-identifying restrictions. With respect to instrument strength, we use the Stock-Yogo (2005) weak identification test (for bias and test size); we also test individual instrument redundancy using a chi-square test. With respect to instrument validity, besides testing over-identifying restrictions for the entire instrument set, we provide tests of the exogeneity of individual (or a subset) of instruments using the C-statistic. Furthermore, to the extent possible, the instrument set used in each regression is not based on the same rationale; this allows us to conduct robustness checks and enhances confidence in the estimates obtained.

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 findings contradict those by Cui *et.al.* (2013), who derived OLS and quantile regression estimates of the return to schooling for migrants as low as 3-5%, concluding that their results raise questions about the 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, not only the return to schooling from OLS for men is slightly higher than for women (which is consistent with past evidence), but the same is the case with IV estimates which are 1-2 percentage points higher for men; these

findings differ from some previous studies which found the opposite (for example, Li, 2005; Chen and Hamori, 2009). Size-wise, we find that the return to an additional year of schooling in urban China, after rising over the transition years, stands at about 10% for men and about 8-9% for women, and is slightly higher for the sample of workers who have migrated to urban centres. Thus, returns to education in urban China in 2009 are of similar magnitude to those for other transition countries (for example, Vietnam), as well as to worldwide and developing country averages. These estimates make more intuitive sense compared to some previous IV estimates of the return to schooling in China, which are in the order of 15-20% per additional year of schooling.

## **6. Conclusion**

This paper contributes to the returns to education in China and instrumental variable (IV) estimation literature by utilizing a new rich dataset, new instrument set and 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 labor markets, are substantial at about 10% for men and 8-9% for women. Male returns are slightly higher compared to female returns, as are returns for Chinese workers who have migrated to major urban centres. The instrumental variable 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. All our estimates (for men, women, urban workers and migrant workers) are within a relatively narrow band, as shown using different instrument combinations in robustness checks. The results suggest that returns to education in urban China have risen over time to levels comparable to those in other transition countries and the 10% world and Asian region averages.

## 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.

Arabsheibani, R. and A. Mussurov (2007). “Returns to Schooling in Kazakhstan.” *Economics of Transition* 15 (2): 341-364.

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

Belmont, L. and F. A. Marolla (1973). Birth Order, Family Size and Intelligence. *Science* 182: 1096-1101.

Bedard, K. and E. Dhuey (2006). “The Persistence of Early Childhood Maturity: international Evidence of of Long-Run Age Effects.” *Quarterly Journal of Economics* 121: 1437-1472.

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, Centre for Economic Policy Research, Australian National University, Canberra.

Bu, Feifei (2014). “Sibling Configurations, Educational Aspiration and Attainment.” Discussion Paper No. 2014-11, Institute for Social and Economic Research, University of Essex.

Chen, G. and S. Hamori (2008). “Economic returns to schooling in urban China: OLS and the instrumental variables approach.” *China Economic Review* 20: 143-152.

Cui, Y., D. Nahm, and M. Tani (2013). Earnings Differentials and Returns to Education in China, 1995-2008. IZA Discussion Paper No. 7349, Institute for the Study of Labour, Bonn, Germany.

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).

Evans, W. N. and E. Montgomery (1994). Education and Health: Where there is Smoke, there is an Instrument. Working Paper 4949, National Bureau of Economic Research, Cambridge, MA.

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.

Fersterer, J. and R. Winter-Ebmer (2003). *Economics of Education Review* 22: 561-566.

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

Fleisher, B. M. and X. Wang (2005). Returns to Schooling in China Under Planning and Reform. *Journal of Comparative Economics* 33: 265-277.

Harmon, C., H. Oosterbeek, and I. Walker (2000). The Returns to Education: A Review of Evidence, Issues and Deficiencies in the Literature. Centre for the Economics of Education, London School of Economics and Political Science, London.

Hesketh T, L. Lu, and Z. W. Xing (2005). The Effect of China's One-Child Policy After 25 Years. *The New England Journal of Medicine* 353: 1171-1176.

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.

Johnson, E. and G. Chow (1997). Rates of Return to Schooling in China. *Pacific Economic Review* 2: 101-113.

Lall, A. and C. Sakellariou (2010). The Evolution of Education Premiums in Cambodia: 1997-2009." *Asian Economic Journal* 24 (4): 333-354.

Li, H. (2003). "Economic Transition and Returns to Education in China." *Economics of Education Review* 22 (3): 317-328.

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

Liu, Z. (1998). "Earnings, Education and Economic Reforms in Urban China." *Economic Development and Cultural Change* 46 (4): 697-725.

Lu, Z. and S. Song (2006). Rural-Urban Migration and Wage Determination: The Case of Tianjin, China. *China Economic Review* 17: 337-345.

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

Meng, X. and M. P. Kidd (1997). "Labor Market Reform and the Changing Structure of Wage Determination in China's State Sector in During the 1980s." *Journal of Comparative Economics* 25 (3): 4-3-421.

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

Murray, M. P. (2006). "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4): 111-132.

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. Institute for the Study of Labour, Bonn, Germany.

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

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):599-510.

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* ( Donald W. K. Andrews and James H. Stock, eds.), Cambridge University Press, pp.80-108.

Strøm B. (2004). Student Achievements and Birthday Effects. Mimeo, Norwegian University of Science and Technology, Trondheim.

Trostel, P., I. Walker, and P. Woolley (2002). Estimates of Economic Return to Schooling for 28 Countries." *Labour Economics* 9 (1): 1-16.

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.

World Development Indicators (2009). The World Bank, Washington, D.C.

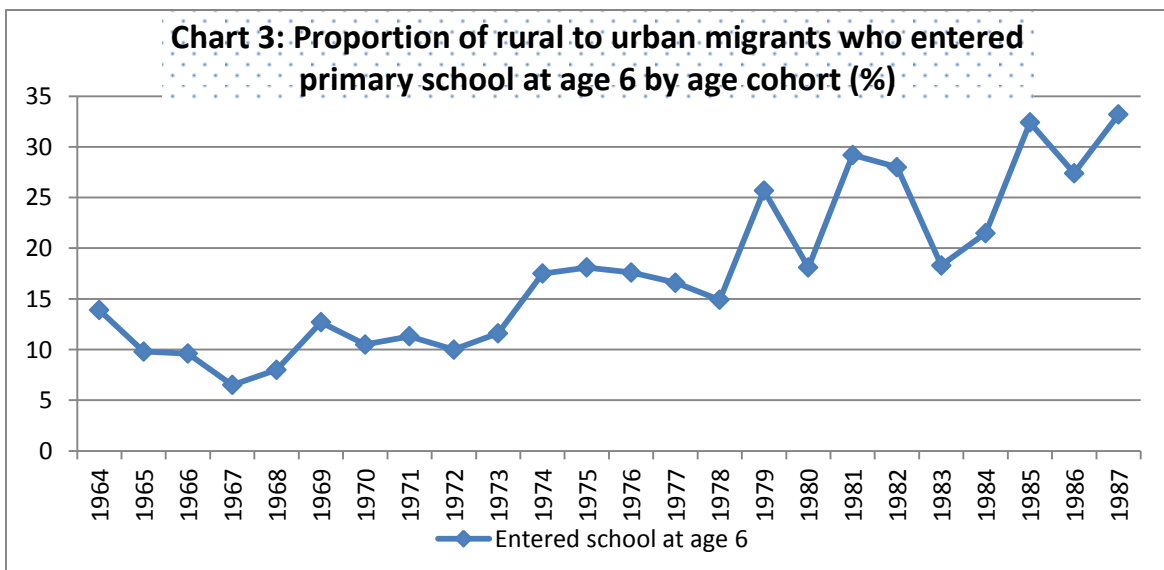
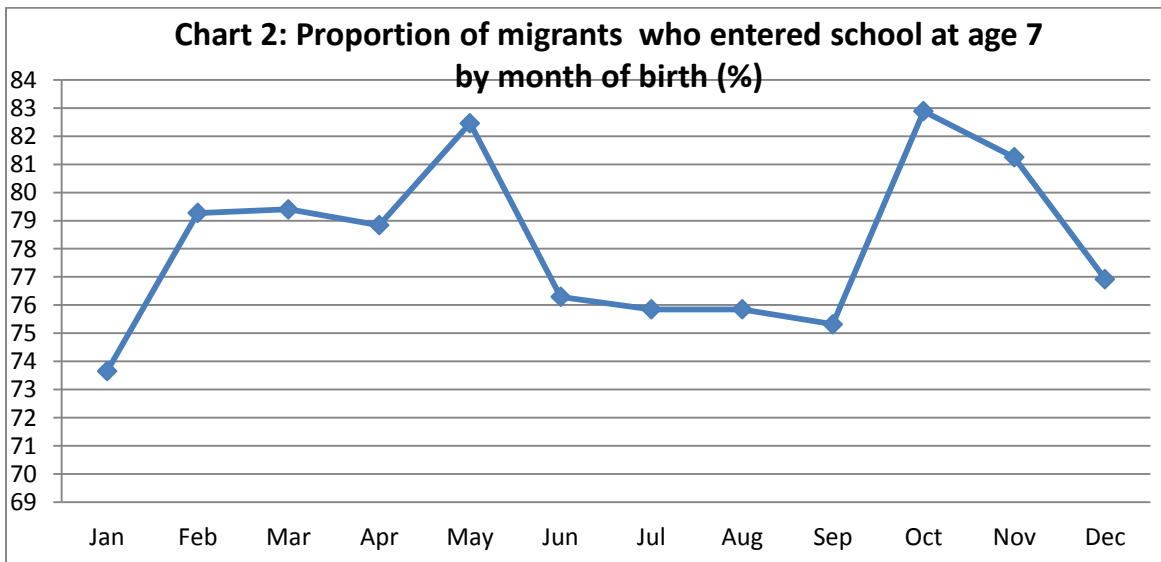
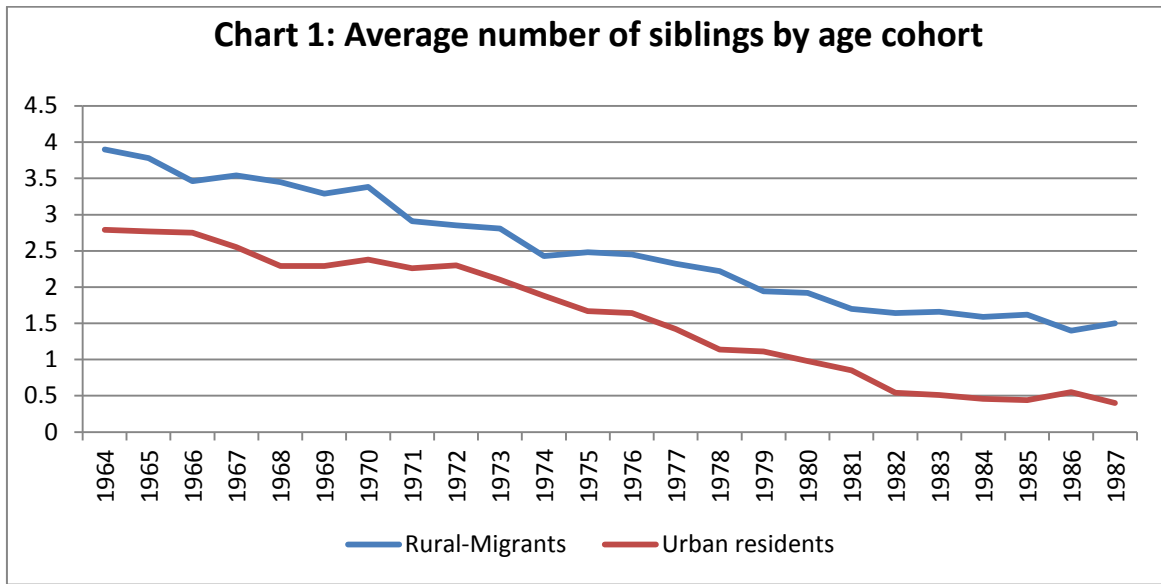
Zajonc, R. B. and G. B. Markus (1975). Birth order and intellectual development. *Psychological Review* 82: 74-88.

Zajonc, R. B. (1976). "Family Configuration and Intelligence." *Science* 192: 227-236.

Zajonc, R. B. and P. R. Mullanly (1997). Birth order: Reconciling conflicting effects. *American Psychologist* 52: 685-699.

Zhang, J., Y. Zhao, P. Albert, and X. Song (2005). Economic Returns to Schooling in Urban China, 1988 to 2001. *Journal of Comparative Economics* 33: 730-752.

## Appendix



Married male urban workers, age 22-45

	OLS <sup>1</sup>	IV <sup>2</sup>	IV <sup>3</sup>	IV <sup>4</sup>
Years of schooling	0.059*** (0.009)	0.089*** (0.017)	0.085*** (0.016)	0.094*** (0.030)
Tenure	0.035** (0.014)	0.045*** (0.013)	0.044*** (0.014)	0.046*** (0.016)
Tenure squared	-0.0013*** (0.0004)	-0.0013*** (0.0004)	-0.0013*** (0.0004)	-0.0013*** (0.0004)
Other experience	-0.011 (0.011)	0.026** (0.013)	0.024* (0.013)	0.029* (0.018)
Other experience squared	-0.0013** (0.0004)	-0.0014*** (0.0004)	-0.0015*** (0.0005)	-0.0015*** (0.0005)
Firm size: 6-20	0.256** (0.100)	0.195** (0.189)	0.223** (0.092)	0.201** (0.097)
Firm size: 21-100	0.302*** (0.089)	0.235*** (0.082)	0.257*** (0.084)	0.242** (0.096)
Firm size: > 100	0.228*** (0.085)	0.181** (0.077)	0.189** (0.078)	0.167* (0.088)
Constant	1.61*** (0.199)	1.10*** (0.314)	1.16*** (0.305)	0.991* (0.535)
<u>First stage:</u>				
Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>		0.268	0.319	0.086
F-value [p-value]		378.7 [0.000]	109.9 [0.000]	30.42 [0.000]
<u>Second stage</u>				
<i>Over-identification test: all instruments:</i>				
Sargan statistic [p-value]		-	0.406 [0.939]	0.457 [0.795]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>				
- Spouse's years of schooling		-	0.001 [0.973]	-
- Early smoker		-	0.292 [0.589]	0.458 [0.499]
- Firstborn		-	0.063 [0.802]	0.312 [0.576]
- Firstborn interacted with age		-	0.046 [0.830]	0.442 [0.506]
<i>Weak identification test</i>				
Cragg-Donald Wald statistic (F-statistic):		378.7	109.9	30.42
Stock-Yogo weak ID test critical values:				
5% maximal IV relative bias		-	16.85	13.91
10% maximal IV size		16.38	24.58	22.20
<i>Redundancy test (Chi=sq. [P=value]) for:</i>				
- Spouse's years of schooling		-	238.1 [0.000]	-
- Early smoker		-	11.54 [0.001]	7.87 [0.005]
- Firstborn		-	59.53 [0.000]	75.43 [0.000]
- Firstborn interacted with age		-	59.65 [0.000]	77.36 [0.000]
<i>Endogeneity test for years of schooling</i>				
Chi-sq. [P-value]		3.35 [0.067]	3.93 [0.048]	1.37 [0.242]
R <sup>2</sup>	0.197			
N	950	1,041	950	985

<sup>1</sup>OLS results using the same sample as for IV estimation in column 3. <sup>2</sup>Using spouse's years of schooling as instrument. <sup>3</sup>Instruments: spouse's years of schooling, early smoking, firstborn and interaction with age. <sup>4</sup> Instruments: Early smoker, firstborn and interaction with age.

Married female urban workers, age 22-45

	OLS <sup>1</sup>	IV <sup>2</sup>	IV <sup>3</sup>	IV <sup>4</sup>
Years of schooling	0.051*** (0.009)	0.086** (0.016)	0.076*** (0.015)	0.075** (0.030)
Tenure	0.011 (0.012)	0.025** (0.012)	0.016 (0.012)	0.018 (0.013)
Tenure squared	-0.0007* (0.0004)	-0.0008** (0.0004)	-0.0007* (0.0004)	-0.0008** (0.0004)
Other experience	-0.033*** (0.011)	-0.005 (0.012)	-0.021* (0.012)	-0.023 (0.017)
Other experience squared	0.0005 (0.0004)	-0.0001 (0.0004)	0.0004 (0.0004)	0.0005 (0.0004)
Firm size: 6-20	0.114 (0.074)	0.131 (0.132)	0.079 (0.076)	0.093 (0.084)
Firm size: 21-100	0.307*** (0.074)	0.282** (0.129)	0.274*** (0.072)	0.281*** (0.080)
Firm size: > 100	0.376*** (0.070)	0.351*** (0.127)	0.342*** (0.069)	0.344*** (0.077)
Constant	1.65*** (0.181)	0.964*** (0.306)	1.24*** (0.264)	1.25** (0.502)
<u>First stage:</u>				
Shea Partial R <sup>2</sup> /Partial R <sup>2</sup>		0.327	0.357	0.087
F-value [p-value]		420.6 [0.000]	173.7 [0.000]	44.87 [0.000]
<u>Second stage</u>				
<i>Over-identification test: all instruments:</i>				
Sargan statistic [p-value]		-	0.349 [0.840]	2.14 [0.143]
<i>Exogeneity/orthogonality of suspect instruments (C-test [P-value]) for:</i>				
- Spouse's years of schooling		-	0.186 [0.667]	-
- Firstborn		-	0.410 [0.522]	-
- Firstborn interacted with age		-	0.176 [0.675]	-
<i>Weak identification test</i>				
Cragg-Donald Wald statistic (F-statistic):		420.6	173.7	44.87
Stock-Yogo weak ID test critical values:				
5% maximal IV relative bias		-	13.91	-
10% maximal IV size		16.38	22.30	19.93
<i>Redundancy test (Chi=sq. [P=value] for:</i>				
- Spouse's years of schooling		-	239.4 [0.000]	-
- Firstborn		-	54.59 [0.000]	72.03 [0.000]
- Firstborn interacted with age		-	58.15 [0.000]	78.88 [0.000]
<i>Endogeneity test for years of schooling</i>				
Chi-sq. [P-value]		4.68 [0.031]	4.33 [0.037]	0.831 [0.362]
R <sup>2</sup>	0.245			
N	948	873	948	951

<sup>1</sup>OLS results using the same sample as for IV estimation in column 4. <sup>2</sup>Spouse's years of schooling as instrument. <sup>3</sup>Instruments: Spouse's years of schooling, firstborn and its interaction with age. <sup>4</sup>Using firstborn and its interaction with age.