

Sorting, Selection, and Transformation of Return to College Education in China

by Belton M Fleisher, Ohio State University,
Haizheng Li, Georgia Tech,
Shi Li, Chinese Academy of Social Sciences
and Xiaojun Wang, University of Hawaii-Manoa

Working Paper No. 05-07

March 2005

Abstract

We estimate selection and sorting effects on the evolution of the private return to schooling for college graduates during China's reform between 1988 and 2002. We find evidence of substantial sorting gains under the traditional system, but gains have diminished and even become negative in the most recent data. We take this as evidence consistent with the growing influence of private financial constraints on decisions to attend college as tuition costs have risen and the relative importance of government subsidies to higher education has declined.

JEL Classification: J31, J24, O15.

Keywords: Return to schooling, sorting gains, heterogeneity, financial constraints, comparative advantage, China.

Corresponding Author: Belton M. Fleisher, Department of Economics, Ohio State University, 1945 N Hight Street, Columbus, OH 43210. <fleisher.1@osu.edu>

Xiaojun Wang, Department of Economics, University of Hawaii-Manoa, 2424 Maile Way, Saunders Hall Room 542, Honolulu, Hawaii, U.S.A. 96822. Phone: (808) 956-7721 <xiaojun@hawaii.edu>

1. Introduction and Background

From the inception of economic reform in China into the early 1990s, wage differences by level of skill, occupation, and/or schooling remained very narrow. The Mincerian return to higher education was quite low in comparison with that in the early years of the Mao era. It was also low relative to other industrialized and industrializing countries including those in several transition economies (Fleisher, Sabirianova, and Wang, 2005). Fleisher and Wang (2005) show that the time path of the return to college education paralleled that to schooling in general. Moreover, college graduates appear to have been severely underpaid relative to their contribution to production (Fleisher and Wang, 2004). There is evidence that in the past 15 years, returns to schooling in China have begun to increase (Zhang and Zhao, 2002; Li, 2003, Yang, 2005). Although rising return to schooling has probably contributed to growing income inequality, it is our view that access to education is a more important factor. According to Yang (1999), China in the late 1990s surpassed almost all countries in the world for which data are available in rising income inequality, and by the year 2000 China found itself with one of the highest degrees of income inequality in the world (Yang, 2002).

We are concerned with the question of whether rising inequality in China is associated with access to educational opportunities. The end of the Mao era saw the influence of political considerations on access to higher education sharply diminish, and college admission criteria reverted to historical practice which placed a very heavy weight on merit as determined by critical tests in junior- and senior high schools. More recently, however, a growing proportion of college students must fund their own educational expenses (Hannum, 2004; Heckman, 2004). The proportion of the population privileged to attend college has been and remains very small by almost any standard, despite a sharp acceleration of schooling expenditures in the past decade (Fleisher and Wang, 2005; Heckman, 2005). The proportion of the population aged 20 and above with a college degree was less than 3.2% in 1993 and grew to 3.5% in 2000 according to the 1993 and 2000 population censuses, respectively (National Bureau of Statistics of China, 1994 and 2002).

Access to college and concomitant economic gain depends not only on current financial resources, but also on the ability to achieve high test scores and on cognitive and other attributes produced in earlier family and educational contexts. Thus, higher educational attainment depends recursively on earlier access to publicly and privately supported education at lower levels as well as on the capacity to borrow funds from family and other sources to pay direct and indirect college costs (Carneiro and Heckman, 2002; Hannum, 2004). If access to all levels of schooling is available only to the financially, politically, and geographically advantaged, the bulk of China's population will be excluded from full participation in the growth of human capital and the income it produces.

In this paper we focus on the returns to college education in China from the end of the first decade of transition to 2002, paying particular attention to sorting and selection issues. We address the following questions.

1. How have the relative importances of variables that determine the probability of college attendance changed?
2. Is there evidence that the degree of sorting into college according to net benefit has changed in the reform period?
 - Has the sorting gain narrowed or widened?
 - If it has widened, is this because more able students are now able to attend college due to reduced favoritism.?
 - If it has narrowed, is this consistent with efficient sorting with an increased proportion of qualified college graduates graduating from college?
 - Is there evidence of increased influence of borrowing constraints on college attendance? (Carneiro and Heckman, 2002)?

2. Methodology

Our method takes into account both heterogeneous returns to schooling and self-selection based on anticipated returns. We first estimate the marginal treatment effect

(MTE) in the sample, which is the building block of other parameters of interest. The marginal treatment effect and its derivatives are estimated using the method developed in Heckman, Ichimura, Todd, and Smith (1998).²

We set up the following model of wage determination by schooling choice:

$$\begin{aligned}\ln Y_1 &= \mu_1(X, U_1) \\ \ln Y_0 &= \mu_0(X, U_0)\end{aligned}$$

where a subscript indicates whether the individual is in the schooled state (1) or the unschooled state (0). Y is income, X is observed heterogeneity, and U is unobserved heterogeneity in wage determination. In general, the functional forms can have a nonlinear component, and $U_1 \neq U_0$.

The schooling choice comes from the following latent dependent model:

$$\begin{aligned}S^* &= \mu_s(Z) - U_s \\ S &= 1 \text{ if } S^* \geq 0\end{aligned}$$

where S^* is a latent variable whose value is determined by an observable component $\mu_s(Z)$ and a unobservable component U_s . A respondent will only attend college (i.e. $S=1$) if this latent variable turns out to be nonnegative.

In our empirical work, Z is a vector of variables that help predict the probability of attending college. It includes parental education, parental income, number of children (siblings), gender, ethnic group, and birth year dummies. On the other hand, X is a vector that holds explanatory power on wages. In the benchmark setting, this includes work experience, work experience squared, gender, ethnic group, ownership, industry, and location. Z and X can share some common variables, but Z must also possess unique variables for the model to be identified.

In the first step, a probit model is used to estimate the $\mu_s(Z)$ function. The predicted value is called propensity score, \hat{P}_i , where the subscript i denotes each individual. The second step adopts a semi-parametric procedure in which *local linear*

regressions are used. Fan (1992, 1993)³ develops the distribution theory for the local linear estimator of $E(Y|P=P_0)$, where Y and P are random variables. They show that $E(Y|P=P_0)$ and its derivatives can be consistently estimated by the following algorithm:

$$\min_{\gamma_1, \gamma_2} \sum_{i \leq N} [Y_i - \gamma_1 - \gamma_2(P_i - P_0)]^2 G\left(\frac{P_0 - P_i}{a_N}\right)$$

where γ_1 is a consistent estimator of $E(Y|P=P_0)$, and γ_2 is a consistent estimator of $\partial E(Y|P=P_0)/\partial P$. $G(\cdot)$ is a kernel function and a_N is the bandwidth. We use a Gaussian kernel and a bandwidth of 0.2 in the estimation.⁴ Obviously, this algorithm is equivalent to applying weighted least square at each observation point, only using samples in its nearest “neighborhood”.

We first estimate $E(\ln Y|P)$ and $E(X|P)$ with the above procedure. Then we run the *double residual regression* of $\ln Y - E(\ln Y|P)$ on $X - E(X|P)$. This is a simple OLS regression, except we trimmed off the smallest 2% of the estimated propensity scores with a biweight kernel as suggested by Heckman, Ichimura, Todd, and Smith (1998). The result is consistently estimated coefficients of the linear components of the model, β .

Define the nonlinear component residual as $U = \ln Y - \beta X$. Use local linear regression again to estimate $E(U|P)$ and its first derivative. This first derivative is the marginal treatment effect (MTE). The average treatment effect (ATE) is a simple integration of the MTE with equal weight assigned to each $P(Z)=U_s$. However, treatment on the treated (TT) and treatment on the untreated (TUT) are calculated with the following weighting functions:

$$h_{TT}(u_s) = \frac{\left[\int_{u_s}^1 f(p) dp \right]}{E(p)}$$

$$h_{TUT}(u_s) = \frac{\left[\int_0^{u_s} f(p) dp \right]}{E(1-p)}$$

where $f(p)$ is the conditional density of propensity scores. The conditioning on X is implicit in the above functions. All integrations are conducted numerically using simple trapezoidal rules.

3. Data

The data used in this study are from the first, second, and third waves of the Chinese Household Income Project (CHIP) conducted in 1989 (CHIP-88), 1996 (CHIP-95), and 2003 (CHIP-2002). We briefly describe our use of the CHIP 95 data here. The data are taken from the urban component of the survey, in which 6,928 households and 21,688 individuals in urban areas of eleven provinces were surveyed for 1995. The survey was funded by the Ford Foundation and a number of other institutes.⁵ In the data, annual earnings include regular wages, bonuses, overtime wages, in-kind wages, and other income from the work unit. The hourly wage rate is calculated based on the reported number of working hours. The education measure includes seven degree categories, ranging from below elementary school to college. For more details, see Li (2003).

In China, the definition of labor force is limited to ages 16 or above. As a general rule, in the late 1970s, children entered elementary school at age 7 and remained there for 5 years; junior high school and senior high school each required 2 years. Thus, an individual who was born in 1962 and started school at age 7 would be a senior in upper middle school in 1978 and face the choice of going to college or starting to work. We limit all of our samples to individuals born after 1961 in order to avoid the complicating effects of educational policy during the Cultural Revolution, when many youths were sent to the countryside for “rectification” (or “re-education”), and colleges and even middle schools were either closed or nonfunctioning. The upper birth-year cutoff eliminates observations born too late to have entered college in China’s education system (for the probit equations) and too late to have completed college (for the wage equations).

Another sample limitation is based on our need for family background information such as parental education and parental income. Thus, our sample is restricted to working individuals who are living in a household with their parents (for the probit equations) and who have positive earnings in 1995 (for the wage equations). As

specified in the model, we only include two education groups: 3 or 4-year college graduates and upper middle school.

4. Empirical Results

First we report empirical analysis of the propensity to attend college and next derive estimates of various treatment effects for college attendance.

4.1 Propensity to Acquire a College Education

Table 1 presents estimates of the probit model for college attendance in the three sample years, 1988, 1995, and 2002. Heckman and Li's (2004) result for the year 2000 is also included for comparison. These probit equations are used to generate a propensity score for each observation, which is the predicted probability of college attendance. The frequency distribution of propensities to attend college provides a reduced-form picture of increasing college attendance in China.

The regressors in table 1 include can be roughly categorized into variables related to ability formation and those related to the budget constraint, although as suggested above, the financial ability to provide childhood investments in human capital may affect measured ability at older ages. For example parental income not only provides the immediate resource to attend college, but also reflects past expenditures, given that income is highly serially correlated. On the other hand, we assume that when parental income is controlled, parental education is a proxy for an offspring's ability.

The columns (4), (8), (9), and (13) of table 1 report the mean marginal propensities (probabilities) attributable to each independent variable. In our sample years 1988, 1995, and 2002, the effect of parental schooling is highly significant, but it becomes quantitatively smaller over time. The marginal impact of a one-year increase in father's education on the probability of a child attending a 4-year college is 2.1 percentage points in 1988, 3.75 percentage points in 1995, but it drops to only 1.72 percentage points in 2002. The impact of mother's education follows the same time pattern. The impact of parental income on college attendance is also significant in most cases. The marginal impact of 1000 yuan/year in combined parental income increasing the probability of attending college is approximately 1.5 percentage points in 1988, 1 percentage point in 1995 and 0.5 percentage point in 2002. These results suggest that

while parental education (an ability proxy) consistently played an important role in children's college attendance, the influence of parental income has declined quantitatively, although it remains statistically significant. Does this suggest that family financial constraints have become less important in college enrollment?

Figure 1 shows the frequency distribution (kernel-smoothed) of propensity scores for 1988, 1995, and 2002.⁶ For each year the left panel shows the distribution for all observations ($S=1$ and $S=0$), while the right panel shows separate distributions for college attenders and nonattenders. The rightward shift of the combined distributions reflects increasing college enrollment and is consistent with the nearly 80% growth of the proportion of the urban population with education of college and above between 1988 and 1995 and more than 100% growth by 1999, as documented in our data and in other studies as well (for example, Zhang and Zhao 2002, table 4). In 1988, the frequency distribution of nonattenders is supported over a range of propensity scores from approximately zero through nearly 0.6⁷; in 1995, it is supported over the range from approximately zero through 0.9, and by 2002, it is supported over almost the entire range of propensities approaching 1.0. The frequency distribution of attenders is supported over the range of propensities between approximately zero and 0.7 in 1988, between approximately zero and greater than 0.9 in 1995, and from about 0.1 through 1.0 in 2002.

There are some interesting implications of comparing the distributions and their shifts over time. Table 2 shows that in 1988, 20.8% of the sample were college graduates and had a propensity score equal to or greater than 0.324. We define this propensity score to be the cutoff score. In 1988 11.4% of the entire sample had scores higher than this value (yet they didn't go to college). The proportion of nonattenders with propensity score above the cutoff for that year rises to 16.8% in 1995 and to 17.6% in 2002. The other group of misfits, namely agents with propensity scores below the cutoff who nevertheless attended college also grew. The proportion of the sample with below the cutoff score who attended college was 12.6% of the sample in 1988, 17.0% in 1995, and 17.4% in 2002. Both patterns suggest that unobserved heterogeneity increased over the period covered in our study, mostly between 1988 and 1995. The increased heterogeneity could reflect (1) a growing proportion of agents with

unobserved financial constraints and high propensity scores who cannot realize their high potential concerns because they are unable to finance college education or (2) a growing importance of unobserved comparative advantage. If (1) dominates, then we should observe sorting gains diminishing over time; if (2) dominates, then sorting gains should increase.

4.2 College Education and Earnings

Tables 3 and 4 contain the results of OLS, IV, and semi-parametric IV (SPIV) estimation of the effect of college attendance on earnings. Table 3 reports the results of benchmark estimates of wage equations in which no proxy variables for student ability are included as regressors. The benchmark OLS estimates for 1988 and 1995 are commensurate with those reported elsewhere for comparable time periods. They show an upward trend in returns to college education, with acceleration after 1995 (See Fleisher and Wang, 2004, for estimates and a summary of other studies)⁸. The IV estimates of the return to college education (all of which use the propensity score as the instrument for college attendance) are considerably higher than the OLS estimates in the benchmark regressions and indicate a greater acceleration after 1995.

Estimates based on regressions containing a proxy for student ability are reported in Table 4. When either parental education or parental income variables are used to proxy for ability, the OLS estimates of the return to schooling are approximately equal to those reported in Table 3, with the exception of the estimate reported by Heckman and Li (2004), for the year 2000. Our OLS estimated return, with parental income used as an ability proxy, is much higher than Heckman and Li's benchmark OLS estimate. When parental education is used as a proxy for student ability in the IV earnings equations, the estimated return to college education is much higher than the OLS estimates for the years 1988 and 1995, and 2002. However, when parental income is used as a proxy for ability, the IV estimates are approximately the same as the OLS estimates in 1988 and 1995, but higher in 2002 (although much lower than when parental education is the ability proxy)⁹.

We turn now to our estimates of returns to schooling based on SPIV estimation. The distinguishing feature of the SPIV procedure is the capacity to retrieve estimates of the marginal treatment effect (MTE) of college education that allow for unobserved heterogeneity in the return to schooling. Figure 2 depicts the MTE of college education in China for the years 1988, 1995, and 2002. For each year we also compare the MTE from two specifications of the wage equation. Figure 3 places these two MTE curves for each year together so that the effect of including an ability proxy can be seen more clearly. Inclusion of an ability proxy in the local linear regressions simply results in an almost parallel upward shift of the MTE curve. The shape is not affected across the Us dimension.

We consistently find that between 1988 and 2002 the average treatment effect, the return to education for a randomly chosen individual, has increased substantially. For example, in the specification with parental education as ability proxy, the rate of return for four years of college has increased from 86.6% [$100(\exp(0.6239) - 1)$], or 16.9% per year of college to 268.5% [$100(\exp(1.3044) - 1)$], or 38.6% per year of college. However, when this dramatic change is decomposed into treatment on the treated (TT) and treatment on the untreated (TUT), we obtain very surprising results. We regard the TT as *realized* return that is obtained by individuals who actually completed four years of college. This realized return to college graduation actually *decreased* from 422.6% [$100(\exp(1.6530) - 1)$] to 126.3% [$100(\exp(0.8168) - 1)$]. In contrast, TUT, which is the counterfactual return for those who did not attend college, jumped from 42.0% [$100(\exp(0.3510) - 1)$] to 687.7% [$100(\exp(2.064) - 1)$]. This implies that the group with the highest potential return actually does not go to college. This surprising result is reflected in the collapse of sorting gains from +179.9% [$100(\exp(1.0291) - 1)$] in 1988 to -38.6% [$100(\exp(-0.4876) - 1)$] in 2002.

The heterogeneity model postulates that those who attend college do so because they benefit more than those who choose not to attend. It is important to emphasize that this assumption does imply that decisions are made strictly in terms of expected income streams. It is consistent with someone choosing not to attend college because financial or psychic costs are expected to outweigh financial gains (Carneiro, Heckman, and Vytlačil 2003). However, if all financial and psychic costs of college attendance are

reflected in the propensity score, the model implies the MTE function is monotonically negatively sloped and represents a demand for college education in the sense that a decline in the marginal financial cost of college attendance is required to induce greater college attendance, *cet. par.* The MTE curves for 1988 support this hypothesis, but they are inconsistent with it in 1995 and, dramatically so, in 2002. The 1995 MTE curves reach a minimum in the middle of the *Us* range and then curve back up toward larger values of *Us*. The 2002 MTE curves are monotonically *increasing* in *Us*. These shapes are inconsistent with the joint hypothesis that agents' unobserved heterogeneity involves only their comparative advantage in ability to benefit from more schooling. They are consistent with some barrier to college attendance in China other than lack of ability to benefit financially, e.g. psychic costs or unobserved financial barriers (Carneiro, Heckman, and Vytlacil 2004, p. 25).

5. Conclusion

. The three estimation methods, OLS, IV, and SPIV, differ substantially in their estimated levels of return to schooling. All three, however, consistently show a substantial growth in returns to schooling between 1995 and 2002. The SPIV measure of returns that is comparable to that obtained by OLS and IV procedures is ATE. These estimates for 2002 range from a low of approximately 7% per year of college for OLS in regressions including parental education as a regressor to 44.4% in IV estimation and 38% in SPIV estimation. When parental income is included as a regressor, but not parental education, the estimated returns to schooling are 7%, 12.6%, and 10.7%, respectively.

IV and SPIV estimates of the return to college are sensitive to the use of a proxy for ability. When parental income is used as a proxy for ability in the local nonlinear wage regression, IV estimated returns to college were unchanged between 1988 and 1995 but nearly tripled between 1995 and 2002. When parental schooling is used as a proxy for ability, the IV estimates are higher than when parental income is used, decreased between 1988 and 1995, and increased sharply between 1995 and 2002.

When parental education is used as a proxy for ability, the SPIV estimate of heterogeneous return per year of collage for college attenders (TT) falls from 51.2% in 1988 to 17.5% in 1995 and then rises to 22.7% in 2002. The counterfactual return per year of college for those who did not attend (TUT) rises substantially, from 9.2% in 1988 to 13.3% in 1995, and to 67.5% in 2002. Sorting gain declines substantially, becoming negative in 2002. This evidence is consistent with the increasing importance of unmeasured financial constraints on college attendance and is the crux of our continued research.

References

- Björklund, Anders and Moffitt, James (1987). "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Review of Economics and Statistics* 69, 1: 42-49.
- Carneiro, Pedro and Heckman, James J. (2002). "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal* 112 (October): 705-734.
- Carneiro, Pedro, Heckman, James J., and Vytlacil, Edward (2003), Understanding what Instrumental Variables Estimate: Estimating the Average and Marginal Returns to Schooling, working paper, Department of Economics, University of Chicago.
- Fan, J. (1992). "Design-adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87: 998-1004.
- _____ (1993). "Local Linear Regression Smoothers and Their Minimax Efficiencies." *The Annals of Statistics* 21: 196-216.
- Fan, J., and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman & Hall.
- Fleisher, Belton (2005). "Higher Education in China: A Growth Paradox." In Kwan, Yum K. and Yu, Eden S. H. (Eds.), *Critical Issues in China's Growth and Development*. Ashgate Publishing Ltd., Aldershot, UK, 3-21.
- Fleisher, Belton M., and Chen, Jian, (1997) "The Coast-Noncoast Income Gap, Productivity, and Regional Economic Policy in China." *Journal of Comparative Economics* 25, 2: 220-36.
- Fleisher, Belton, Dong, Keyong, and Liu, Yunhua, (1996). "Education, Enterprise Organization, and Productivity in the Chinese Paper Industry." *Economic Development and Cultural Change* 44, 3(April): 571-587.
- Fleisher, Belton, Sabirianova, Klara, and Wang, Xiaojun (2005). "Returns to Skills and the Speed of Reforms: Evidence from Central and Eastern Europe, China, and Russia." *Journal of Comparative Economics* 33, 2 (in press).
- Fleisher, Belton M. and Wang, Xiaojun (2001) "Efficiency Wages and Work Incentives in Urban and Rural China." *Journal of Comparative Economics* 29, 4: 645-662.

-
- (2004). "Skill Differentials, Return to Schooling, and Market Segmentation in a Transition Economy: The Case of Mainland China." *Journal of Development Economics* 73, 1: 715-728.
- Fleisher, Belton and Wang, Xiaojun (2005). "Returns to Schooling in China under Planning and Reform." *Journal of Comparative Economics* 33, 2 (in press).
- Fleisher, Belton M. and Yang, Dennis Tao. (2003). "China's Labor Markets," working paper, Department of Economics, The Ohio State University
- Giles, John, Park, Albert, and Zhang, Juwei (2004). "The Great Proletarian Cultural Revolution, Disruption to Education, and Returns to Schooling." East Lansing, Michigan: Department of Economics, Michigan State University.
- Hannum, Emily and Wang, Meiyan (2004). "Geography and Educational Inequality in China." Working paper, Department of Sociology and Population Studies Center, University of Pennsylvania.
- Heckman, James J. (2005). "China's Human Capital Investment." *China Economic Review* 16, 1: 50-70.
- Heckman, James, Ichimura, Hidehiko, Todd, Petra, and Smith, Jeffrey (1998). "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66, 5: 1017-1098.
- Heckman, James J. and Li, Xuesong (2004). "Selection Bias, Comparative Advantage, and Heterogeneous Returns to Education: Evidence from China in 2000." *Pacific Economic Review* 9, 155-171.
- Jones, Derek C. and K. Ilayperuma (1994). "Wage Determination under Plan and Early Transition: Evidence from Bulgaria." Working Paper No. 94/7, Hamilton, NY: Department of Economics, Hamilton College.
- Knight, John and Song, Lina (1991). "The Determinants of Urban Income Inequality in China." *Oxford Bulletin of Economics and Statistics* 53, 2: 123-154.
- Li, Haizheng (2003). "Economic Transition and Returns to Education in China." *Economics of Education Review* 2, 317-328.
- Liu, Zhiqiang (1998). "Earnings, Education, and Economic Reform in China." *Economic Development and Cultural Change* 46, 4: 697-726.

- Meng, Xin (2000). *Labour Market Reform in China*. Cambridge, UK: Cambridge University Press.
- Meng, Xin and Gregory, R. G. (2002). "The Impact of Interrupted Education on Subsequent Educational Attainment: A Cost of the Chinese Cultural Revolution." *Economic Development and Cultural Change* 50, 4: 934-955.
- Meng, Xing and Kidd, Michael P (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.
- Munich, Daniel, Svejnar, Jan, and Terrell, Katherine (2000). "Returns to Human Capital under the Communist Wage Grid and During the Transition to a Market Economy." Paper presented at the World Conference of the European Association of Labor Economists and Society of Labor Economists, Milan, Italy.
- National Bureau of Statistics of China (1994), Annual Population Change Survey 1993, Beijing, China.
- National Bureau of Statistics of China (2002), 2000 Population Census, Beijing, China.
- Orazem, Peter and Vodopivec, Milan (1995). "Winners and losers in the Transition: Returns to Education, Experience, and Gender in Slovenia." *The World Bank Review* 9, 2: 201-230.
- Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.
- Yang, Dennis Tao (1999). "Urban-Biased Policies and Rising Income Inequality in China." *American Economic Review* 89, 2: 306-310.
- Yang, Dennis Tao (2002). "What Has Caused Regional Inequality in China?" *China Economic Review* 13, 4: 331-334.
- Yang, Dennis Tao (2005). "Determinants of Schooling Returns during Transition: Evidence from Chinese Cities." *Journal of Comparative Economics* 33,2 (in press).
- Zhang, Junsen and Zhao, Yaohui (2002). "Economic Returns to Schooling in Urban China, 1988-1999." Paper presented at the 2002 meetings of the Allied Social Sciences Association, Washington DC.

- Zhou, Xueguang (2000). "Economic Transformation and Income Inequality in Urban China: Evidence from Panel Data." *American Journal of Sociology* 105, 4 (January): 1135-74.
- Zhou, Xueguang and Hou, Liren (1999). "Children of the Cultural Revolution: The State and the Life Course in the People's Republic of China." *American Sociological Review* 65,1 (February): 12-36.
- Zhou, Xueguang and Moen, Phyllis (2002), The State and Life Chances in urban China, 1949-1994 (Computer file). ICPSR version. Durham, NC: Duke University, Dept of Sociology [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Table 1: Propensity Estimates

Variable	CHIP88				CHIP95				H&L(2000)	CHIP02			
	Param.	t-ratios	p-values	(4) Mean Marginal Effect	Param.	t-ratios	p-values	(8) Mean Marginal Effect	(9) Mean Marginal Effect	Param.	t-ratio	p-values	(13) Mean Marginal Effect
CONST	-1.3255	-5.7944	0.0000	-0.3514	-1.9984	-7.9504	0.0000	-0.7868		-0.9850	-4.5055	0.0000	-0.3677
FEDU	0.0802	5.2656	0.0000	0.0213	0.0953	6.2347	0.0000	0.0375	0.0211	0.0462	3.6987	0.0001	0.0172
MEDU	0.0159	1.0198	0.1540	0.0042	0.0718	4.8196	0.0000	0.0283	0.0126	0.0448	3.4419	0.0003	0.0167
FWAGE	0.0302	0.9124	0.1809	0.0080	0.0161	0.6243	0.2663	0.0063	0.0040*	0.0132	2.3574	0.0093	0.0049
MWAGE	0.0889	1.9831	0.0238	0.0236	0.0418	1.2254	0.1104	0.0165		0.0139	1.8970	0.0290	0.0052
CHL	-0.1287	-2.4677	0.0069	-0.0341	-0.0430	-0.6533	0.2569	-0.0169		-0.0471	-0.5598	0.2878	-0.0176
SEX	-0.0082	-0.0967	0.4615	-0.0022	-0.1196	-1.3953	0.0816	-0.0471		-0.3658	-5.2193	0.0000	-0.1366
ETHNIC	-0.0247	-0.1054	0.4580	-0.0066	-0.2063	-0.9342	0.1752	-0.0812		-0.1652	-1.1316	0.1290	-0.0617
BY1968	-0.9831	-6.4701	0.0000	-0.2606	-0.2089	-1.4676	0.0713	-0.0823		0.4368	2.5557	0.0053	0.1631
BY1967	-0.5369	-3.2247	0.0006	-0.1423	0.1256	0.7025	0.2413	0.0495		0.3796	1.4943	0.0676	0.1417
BY1966	-0.4494	-2.7560	0.0030	-0.1191	0.0507	0.2882	0.3866	0.0200		0.4366	1.5641	0.0590	0.1630
BY1965	-0.4114	-2.5360	0.0057	-0.1091	0.0320	0.1823	0.4277	0.0126		0.2397	0.8472	0.1985	0.0895
BY1964	-0.2138	-1.4434	0.0746	-0.0567	0.0489	0.2662	0.3951	0.0192		0.4112	1.3665	0.0860	0.1535
BY1963	-0.2327	-1.5030	0.0665	-0.0617	0.0479	0.2253	0.4109	0.0188		-0.1613	-0.4929	0.3111	-0.0602
BY1962					0.3006	1.3692	0.0856	0.1182		0.2677	0.8435	0.1996	0.0999

Notes: The dependent variables is a dummy variable = 1 for graduated from 3- or 4-year college. The independent variables are, respectively, father's education in years, mother's education in years, mother's and father's annual income in 1000 *yuan* per year, including cash and in-kind benefits, number of children in family of origin, a dummy variable = 1 if respondent is male, dummy variable =1 if ethnicity is not Han Chinese, and dummy variables for birth year.

*The coefficient is for the variable parental income.

Table 2: Comparison of Propensity Distributions

	1988	1995	2002
Proportion of sample who are college attenders or graduates	20.8%	42.7%	63.0%
Cutoff Propensity	0.324	0.480	0.588
Number of nonattenders with scores above the cutoff as proportion of sample	11.4%	16.8%	17.6%
Number of attenders with scores below the cutoff as proportion of sample	12.6%	17.0%	17.4%

Notes: The cutoff percentage is the propensity score that corresponds to the cumulative frequency of the total sample that were attending or had graduated from college in the sample year.

Table 3: Benchmark regression estimates and treatment effect estimates

Parameter	CHIP88	CHIP95	CHIP02	H&L (2000)
OLS	0.1986 (0.0341)	0.2307 (0.0441)	0.3142	0.0856
IV	0.3435 (0.1063)	0.3724 (0.1216)	0.9812	0.2192
ATE	0.2556 (0.0932)	0.3473 (0.2129)	0.8248	0.2321
TT	0.7868 (0.2577)	0.3883 (0.3650)	0.3943	0.1909
TUT	0.1147 (0.1008)	0.3135 (0.2525)	1.4958	0.2679
Bias = OLS - ATE	-0.0569 (0.0978)	-0.1166 (0.2186)	-0.5106	-0.1465
Selection Bias = OLS - TT	-0.5882 (0.2539)	-0.1576 (0.3712)	-0.0800	-0.1053
Sorting Gain = TT - ATE	0.5312 (0.2260)	0.0410 (0.2477)	-0.4306	-0.0412
TT - TUT	0.672	0.0748	-1.102	-0.077

Notes: Dependent variable is monthly wage in 1988, hourly wage in 1995 and 2002. OLS regressors are a dummy variable for college attendance, experience, experience squared, a dummy variable = 1 if male, a dummy variable = 1 if ethnicity not Han Chinese, The IV regression uses predicted college attendance based on the propensity score as an instrument. The treatment effect estimates are based on results from local linear regression. Standard errors are obtained by bootstrapping. All coefficients represent the estimated return to four years of college.

Table 4: Regression estimates with ability proxy included and treatment effect estimates

Parameter	CHIP88	CHIP88	CHIP95	CHIP95	CHIP02	CHIP02	H&L (2000)
Ability proxy	fedu medu	fwage mwage	fedu medu	fwage mwage	fedu medu	fwage mwage	parental income
OLS	0.2029 (0.0374)	0.1985 (0.0376)	0.2127 (0.0512)	0.2114 (0.0516)	0.2814	0.2687	0.2929
IV	0.8494 (0.2213)	0.2033 (0.1093)	0.5963 (0.5084)	0.1995 (0.1437)	1.4711	0.4764	0.5609
ATE	0.6239 (0.1596)	0.1854 (0.1111)	0.5660 (0.5139)	0.1889 (0.2538)	1.3044	0.4084	0.4336
TT	1.6530 (0.3888)	0.5817 (0.2506)	0.6460 (0.7096)	0.2215 (0.3735)	0.8168	0.2025	0.5149
TUT	0.3510 (0.1436)	0.0804 (0.1183)	0.5002 (0.4284)	0.1621 (0.2788)	2.064	0.7293	0.3630
Bias = OLS - ATE	-0.4211 (0.1460)	0.0130 (0.1056)	-0.3533 (0.5041)	0.0226 (0.2360)	-1.023	-0.1397	-0.1407
Selection Bias = OLS - TT	-1.4502 (0.3717)	-0.3832 (0.2460)	-0.4333 (0.6990)	-0.0100 (0.3583)	-0.5354	0.0662	-0.2220
Sorting Gain = TT - ATE	1.0291 (0.2972)	0.3963 (0.2142)	0.0800 (0.2797)	0.0326 (0.2265)	-0.4876	-0.2059	0.0813
TT - TUT	1.302	0.5013	0.146	0.0594	-1.25	-0.527	0.155

Notes: Dependent variable is monthly wage in 1988, hourly wage in 1995 and 2002. OLS regressors are a dummy variable for college attendance, experience, experience squared, a dummy variable = 1 if male, a dummy variable = 1 if ethnicity not Han Chinese. The IV regression uses predicted college attendance based on the propensity score as an instrument. The treatment effect estimates are based on results from local linear regression. Standard errors are obtained by bootstrapping.

**Figure 1: Propensity to Attend College
(Frequency Distributions of 1988, 1995 and 2002)**

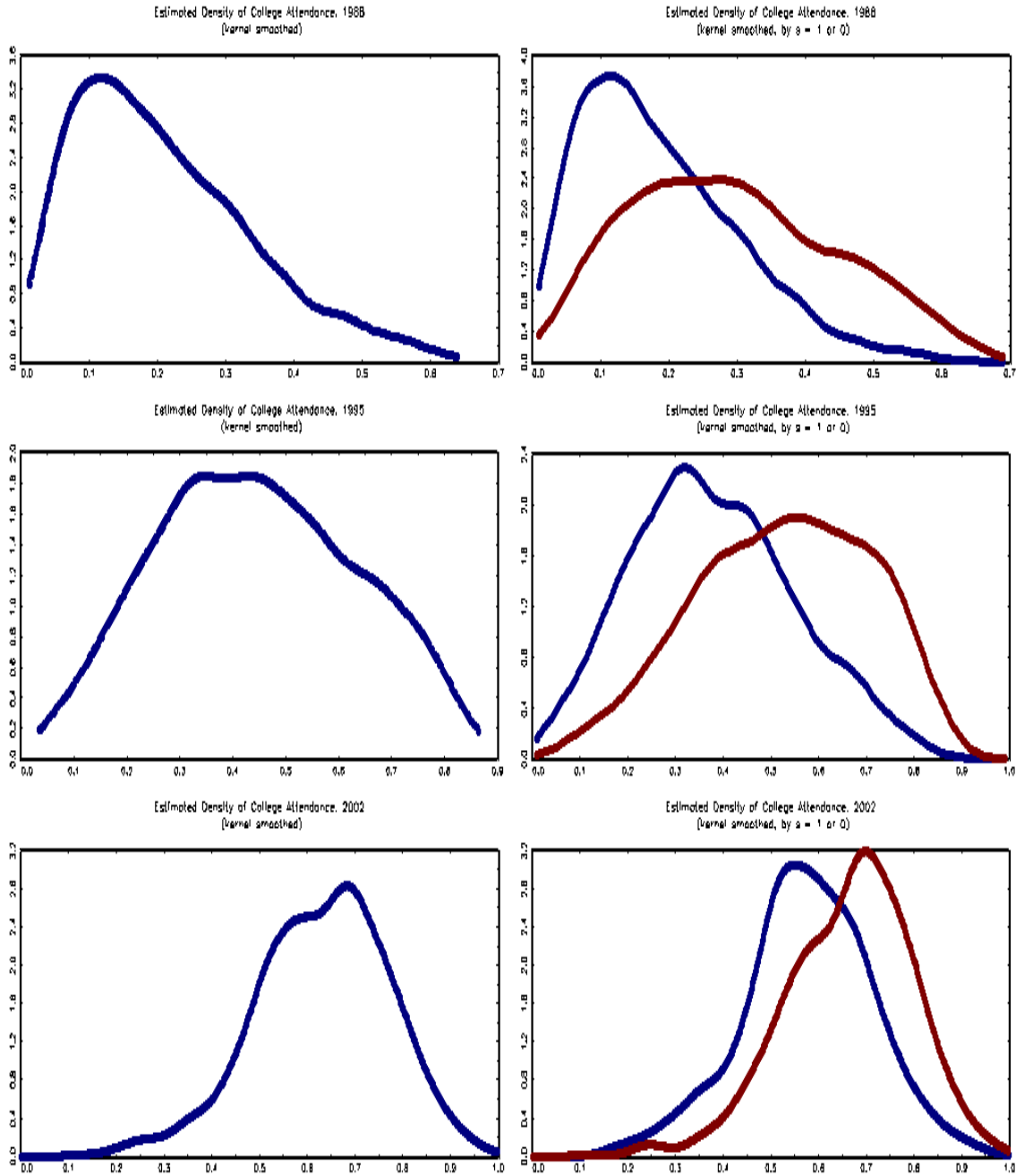
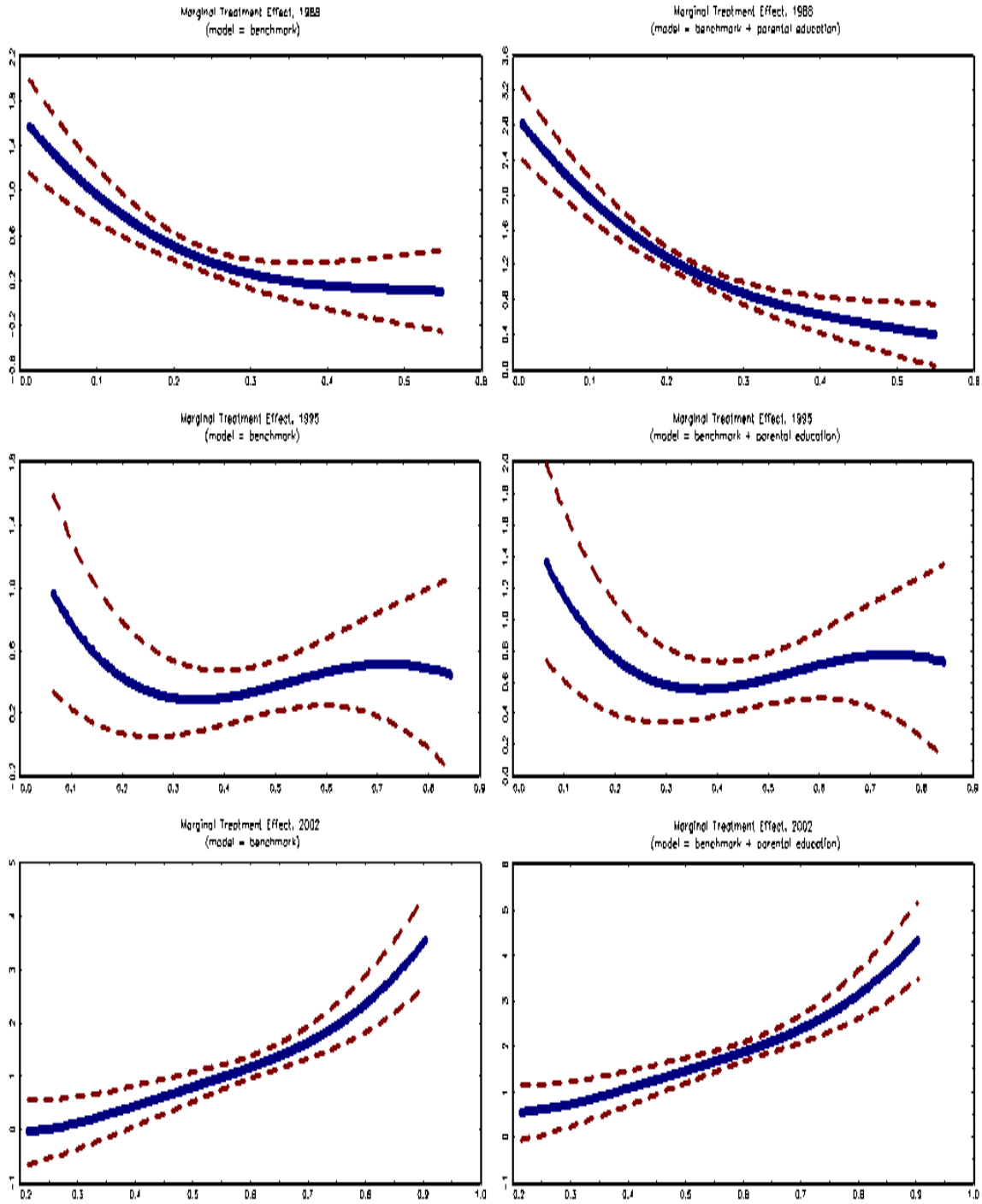
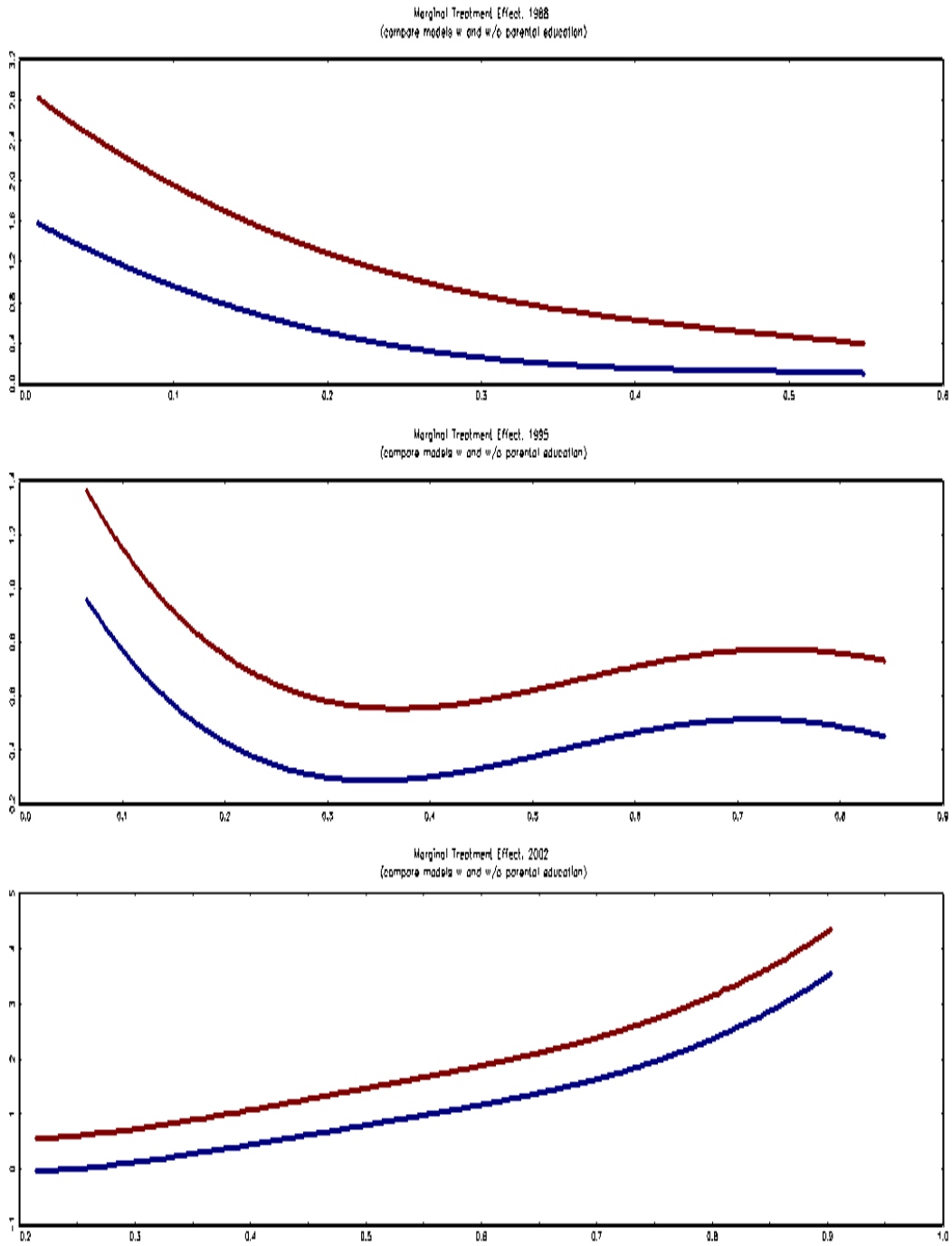


Figure 2: Marginal Treatment Effects: 1988, 1995 and 2002
 (Red dashed line: one standard error bound)



**Figure 3: MTE Curves with and without ability proxies
(parental education, red/upper line)**



¹ We are grateful to Pedro Carneiro, Joe Kaboski, and James Heckman for their invaluable help and advice and to Sergio Urzúa for providing help and advice with software codes. Quheng Deng contributed invaluable research assistance.

* Corresponding author.

² These derivatives include average treatment effect (ATE), treatment on the treated (TT), treatment on the untreated (TUT), bias, selection bias, and sorting gain.

³ Fan (1992): *Journal of the American Statistical Association* 87: 998-1004. Fan (1993): *The Annals of Statistics* 21: 196-216.

⁴ This approximates the rule-of-thumb bandwidth selector proposed in Fan and Gilbels (1996).

⁵ The CHIP-95 data are available to the public at the Inter-university Consortium for Political and Social Research (ICPSR).

⁶ The sample densities are smoothed with Gaussian kernels with optimal bandwidths defined in Silverman (1986).

⁷ A small support implies other factors play important roles, or large unobservable heterogeneity in college education.

⁸ The OLS estimate of return to schooling in 2000 reported by Heckman and Li (2004) is problematic. In their benchmark regression, they report an OLS estimate of 0.0856 for four years of college, implying an annual rate of return of only 2.1%, which is much lower than estimated returns in 1988 and 1995; in an OLS regression that includes parental income as a proxy for ability, they report an estimate of 0.2929 for four years of college, implying an annual rate of return of 6.6%, about the same as in 1988 and 1995. Moreover, the OLS estimates reported by Heckman and Li (2004) are low in comparison to those obtained in other research. Giles, Park, and Zhang (2004) use data for the year 2000 obtained from the China Urban Labor Survey conducted in 2001. The data cover the cities of Fuzhou, Shanghai, Shenyang, Wuhan, and Xian. Using these data, they obtain an estimate for return to four years of college education of

approximately 0.52, which converts to approximately 11% annual rate (personal conversation with John Giles).

⁹ Heckman and Li (2004), however, report an IV estimate of return to schooling equal to 0.5609 for college graduates based on a regression in which parental income is used as a proxy for ability. This is nearly twice as large as their reported OLS estimate.