Access to Higher Education and Inequality: The Chinese Experiment

February 16, 2009

Xiaojun Wang[a, b]
Belton M. Fleisher[c]
Haizheng Li[d]
Shi Li[e]

[b] Department of Economics, University of Hawaii at Manoa, 2424 Maile Way 527, Honolulu, HI 96822, email: xiaojun@hawaii.edu; and China Center for Human Capital and Labor Market Research, Central University of Finance and Economics, Beijing, China.
[c] Corresponding author. Department of Economics, The Ohio State University, 1945 N. High St., Columbus, OH 43210, email: fleisher.1@osu.edu; and China Center for Human Capital and Labor Market Research, Central University of Finance and Economics, Beijing, China; and IZA.
[d] Department of Economics, Georgia Institute of Technology, Atlanta, Georgia; and China Center for Human Capital and Labor Market Research, Central University of Finance and Economics, Beijing, China.
[e] School of Economics and Business, Beijing Normal University, Beijing, China.

Access to Higher Education and Inequality: The Chinese Experiment

Abstract

We apply a semi-parametric latent variable model to estimate selection and sorting effects on the evolution of private returns to schooling for college graduates during China's reform between 1988 and 2002. We find that there were substantial sorting gains under the traditional system, but they have decreased drastically and are negligible in the most recent data. We take this as evidence of growing influence of private financial constraints on decisions to attend college as tuition costs have risen and the relative importance of government subsidies has declined. The main policy implication of our results is that labor and education reform without concomitant capital market reform and government support for the financially disadvantaged exacerbates increases in inequality inherent in elimination of the traditional "wage-grid."

I. Introduction and Background

Two salient features of the labor force in centrally planned economies were the wage-grid and the *nomenklatura*. The wage-grid system compressed wage differentials across education groups, while the *nomenklatura* system selected who attended college to acquire knowledge and training to function in the planning bureaucracy. The private economic return to higher education in terms of earnings tended to be very low. China, since 1978, the former Soviet Union and its satellites, since approximately 1990, have given up central planning and entered a period of transition to market systems. During transition, wage-grids have been relaxed or removed, and wage differentials have increasingly reflected market outcomes; educational attainment, especially at higher levels, has become subject to conscientious choices made by each individual; conventionally estimated returns to education have risen to levels comparable to those observed in developed countries. However, transition toward free markets has occurred at different speeds across the formerly planned economies, and wage differential trajectories have varied widely.[1]

Among the major transitional economies, China has taken the most gradual course toward market reform. From the inception of economic reform into the early 1990s, wage differences by level of skill, occupation, and/or schooling remained very narrow. The Mincerian return to higher education was even lower than in the early years of the Mao era (i.e. 1950s; see Fleisher and Wang 2005). Since the early 1990s, there is evidence that returns to schooling in China have begun to increase (Zhang and Zhao, 2002; Li, 2003; Yang, 2005). Although the rising return to schooling most likely has contributed to growing income inequality,[2] a major concern addressed here, however, is that growing inequality has been exacerbated by in increased difficulty for many families to *access*

---

[1] There is a growing literature on returns to education and wage differentials experienced in these transitional economies. See Brainerd (1998) on Russia; Munich, Svejnar and Terrell (2005) on the Czech Republic; Orazem and Vodopivec (1995) on Slovenia; and Jones and Ilayperuma (2005) on Bulgaria. Fleisher, Sabirianova and Wang (2005) provide a comparative study of eleven former centrally planned economies including Russia and China.

[2] Yang (2005) shows that the dispersion of returns to schooling across Chinese cities increased sharply between 1988 and 1995. Wang et al (2007) provide most recent evidence of rising income inequality from 1987 to 2002 in China.

educational opportunities. Evidence that this concern is well founded is vividly presented by Hannum and Wang (2006). Using 2000 Census data, they show that the percent of variation in years of schooling explained by birth province increased significantly during our sample period. According to Yang (1999, 2002), China in the late 1990s surpassed almost all countries in the world for which data are available in rising income inequality.

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 senior high schools. More recently, however, a growing proportion of college students have had to fund their own educational expenses (Hannum and Wang, 2006; Heckman, 2004), forcing them to forego college due to financial constraints.[3] By 1997, tuition became mandatory in all colleges in China, and the average tuition reached about 31% of per capita GDP. This ratio rose to 46% in 2002, roughly the same level as for private colleges and universities in the US (Li 2009). Between 1992 and 2003, the government share in total education expenditures in China decreased from 84% to 62%, and the share of tuition and fees increased from about 5% to approximately 18% (*China Statistical Yearbook* 2005). 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 and expansion of enrolment in the past decade (Fleisher and Wang, 2005; Heckman, 2005), 0.6% in 1982, 1.4% in 1990, 2.0% in 1995, 4.1% in 2001, and 6.2% in 2006, according to various issues of *China Statistical Yearbook*.

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 to pay direct and indirect college costs (Carneiro and Heckman, 2002; Hannum and Wang, 2006). If access to all levels of schooling is available only to the financially, politically,

---

[3] Until the early 1990s, college education was almost free in China. The government paid for tuition and lodging, while students only needed to pay for meals and books. Tuition at major Chinese Universities now approaches US$1,000 per year or more (People's Daily, 2000).

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 *changes* in returns to college education during the course of economic transition in China from 1988 to 2002. Unlike most traditional literature on this topic that assumes homogenous returns, we apply a semi-parametric estimator, assuming that returns are heterogeneous returns and that individuals respond to anticipated returns. We pay particular attention to sorting, selection, and cohort-specific treatment effects and their changes over time as China marketizes its higher education system. We address the following questions.

1. How do the estimation results based on heterogeneous returns assumption differ from those obtained with traditional OLS and IV methods?
2. How have the determinants of the probability of college attendance changed?
3. Have the gains, or potential gains, from choosing college narrowed or widened? Are the changes caused by self-selection based on comparative advantage or by involuntary selection?
4. Is there evidence that financial constraints have reduced the effectiveness of higher education in China?

Our major contribution is to estimate both the levels of and *changes* in the returns to a four-year college education over a critical time period of China's transition toward free markets. But "free to choose" implies "required to pay" as well. Is the transition to the need for students and their families to finance an increasing proportion of the cost of higher education matched by an equally rapid change in provisions for the financially constrained? The literature has largely ignored the impact of lagging capital market reform on individual investment in human capital. In this paper, we shed some light on the effects of this lack of coordination in reform.

We exploit three cross-sectional data sets, collected in 1988, 1995, and 2002.[4] Our three sample years represent three distinct stages of China's transition from tuition-free college with some living allowances through the 1989 beginnings of the transition away from "free" college education to mandatory tuition in all colleges. Our data allow us to

---

[4] These are not panel data sets.

investigate the impact of these fundamental institutional changes. By 2002, this transition was well advanced. Throughout this period and especially after 1998, higher education capacity expanded rapidly. Universities increased enrollment substantially, and the government initiated a number of policies to foster world-class universities in China.[5]

The large body of literature on returns to education abounds with studies that assume homogenous returns. Following Griliches (1977), a great deal of effort has been devoted to correcting bias caused by unobserved ability and measurement error (Card 1999). However, the instrumental variables (IV) method suggested to correct bias in the estimators breaks down when returns are *heterogeneous.* Another strand of research follows the work pioneered by Roy (1951), Willis and Rosen (1979), and Willis (1986). These scholars assume that schooling decisions are conscientious choices by rational forward-looking individuals who act on their anticipated *heterogeneous* returns to education. Under these conditions, the appropriate procedure is to estimate a latent variable model with correlated random coefficients.

We use methods developed in Heckman and Vytlacil (1999, 2000) that combine the treatment effect literature (Bjorklund and Moffitt, 1987) with the latent variable literature. Heckman and Vytlacil (1999, 2000) and Caneiro, Heckman, and Vytlacil (2000) explain why conventional approaches fail to estimate meaningful policy parameters when agents act on anticipated heterogeneous returns. Suppose the return to schooling β is randomly distributed across the population as shown in Figure 1. Ignoring the heterogeneity and uncertainty in the costs of attaining education, let β_1 be the current breakeven return. That is, only those agents whose return to education is greater than β_1 will find it worthwhile to attend school. There are several interesting policy parameters in this framework, but it is unclear which one the conventional instrumental variable method estimates. For example, the mean return for those who attend school is $\int_{\beta_1}^{\infty} \beta dF(\beta)$ where $F(\beta)$ is the cumulative distribution function of the returns to education, the mean (counterfactual) return for those who do not attend school is $\int_{-\infty}^{\beta_1} \beta dF(\beta)$, and the population mean return is $\int_{-\infty}^{\infty} \beta dF(\beta)$. Suppose a tuition hike

---

pushes the breakeven return up to β_2, then the conventional instrumental variable method estimates $\int_{\beta_1}^{\beta_2} \beta dF(\beta)$ — the average return of those whose schooling decisions are reversed due to the tuition hike, which in general doesn't agree with any of the three parameters described above.[6] That is, the conventional instrumental variable method doesn't recover appropriate policy parameters because the subset of returns of those who reverse decisions due to the instruments is not representative of the schooled, the unschooled, or the population as reflected in the entire hypothetical distribution of returns depicted in Figure 1.

In this paper we estimate cohort-specific parameters that answer well-posed policy questions under the assumption of heterogeneous returns: (i) Average return to college, i.e. the average treatment effect, which measures the return to a randomly selected individual in the sample; (ii) the treatment on the treated effect, which measures the return to those who actually attend college; and (iii) the treatment on the untreated effect, which estimates the potential (counterfactual) gain a high-school graduate could earn had he/she attended college. These estimates reveal important information for individuals' schooling choice and for government's education policies. For instance, such information can help design government assistance to the financially disadvantaged to attend college, or for government policy to expand enrollment so that higher education becomes more accessible.

The rest of the paper is organized as follows. Section II presents the theoretical framework and derives the parameters enumerated above. Section III briefly discusses the data. Empirical results are reported and analyzed in section IV. Section V draws conclusion.

---

[6] The same also applies to other popular instrumental variables used in the literature such as compulsory schooling and distance to nearest schools, etc.

II. Methodology

Our method takes into account both heterogeneous returns to schooling and self-selection based on anticipated returns. We set up the following model of earnings determination by schooling choice:

$$\ln Y_1 = \mu_1(X) + U_1$$
$$\ln Y_0 = \mu_0(X) + U_0$$

$$(1)$$

where a subscript indicates whether the individual is in the schooled state ($S=1$) or the unschooled state ($S=0$).[7] Y is a measure of earnings, and X is observed heterogeneity that might explain earnings differences. U_1 and U_0 are unobserved heterogeneities in earnings determination, and $E(U_0)=E(U_1)=0$. The vector X includes work experience, work experience squared, gender, ethnicity, and occupational characteristics. In general, the functional forms can have a nonlinear component, and $U_1 \neq U_0$.

Each individual can choose only one of the above two states. The schooling choice decision is described by the following latent variable model:

$$S^* = \mu_s(Z) - U_s$$
$$S = 1 \ \ if \ \ S^* \geq 0$$
$$S = 0 \ \ otherwise$$

$$(2)$$

where S^* is a latent variable whose value is determined by an observed component $\mu_s(Z)$ and an unobserved component U_s. A rational individual will attend college (i.e. $S=1$) only if this latent variable is nonnegative. In our empirical work, the vector Z may share some variables with vector X, but Z must also contain variables not in X for the model to be identified. In the vector Z we include parental education, parental income, number of

---

[7] Throughout this paper the schooled state is attending college, while the unschooled state is not attending college after graduating from high school. Sometimes the college state is also referred to as the treated state, while high-school graduates are sometimes referred to as the untreated state. We only consider individuals who at least have graduated from high school.

siblings, gender, ethnicity, and birth year dummies. Variables included in *Z* but not in *X* serve as instruments to identify the returns to education, and these instruments are applied *locally* so that they identify each region in the distribution of the marginal treatment effects (discussed below).[8] Equations (1) and (2) are correlated not only because *X* and *Z* usually share components, but also because the schooling decision at least partially depends on anticipated returns implied in the potential earnings equations and thus the unobservables are also correlated.

In estimating the schooling choice model, we use both parental income and parental education to control for ability formation and for possible financial constraints. Research on human resources is abundant with evidence that children from well-educated parents are more likely to go to college (e.g. Ashenfelter and Zimmerman, 1997). Higher parental income not only mitigates short-run financial constraints, it also predicts long-term ability-enhancing benefits due to better earlier education, better nutrition, and better environments that foster cognitive and non-cognitive skills in children.

From equations (1), we define a heterogeneous return to education,

$$\beta = \ln Y_1 - \ln Y_0 = \left(\mu_1(X) - \mu_0(X)\right) + \left(U_1 - U_0\right) \tag{3}$$

Therefore β is a random variable correlated with U_0 and U_1. Pooling the schooled and unschooled together,

$$\ln Y = S \ln Y_1 + (1-S)\ln Y_0 = \ln Y_0 + \beta S = \mu_0(X) + \beta S + U_0. \tag{4}$$

Equations (3) and (4) reveal the problems in conventional OLS estimation. More specifically, Heckman and Li (2004) shows[9]

---

[8] The conventional *global* instrument method (see Figure 1) only identifies the mean return of the subset of people whose decisions are reversed by the instrument. However this subset does not in general represent the treated, the untreated, or the population. By applying an instrument locally, we circumvent the "representativeness" issue by identifying a limiting version of the return, i.e. the marginal treatment effect.
[9] We suppress the conditioning of *X* here and below in order to simplify exposition.

$$p \lim\left(\hat{\beta}_{OLS}\right) = E\left(\ln Y_1 \mid S = 1\right) - E\left(\ln Y_0 \mid S = 0\right)$$
$$= E\left(\mu_1(X) - \mu_0(X)\right) + \left[E\left(U_1 \mid S = 1\right) - E\left(U_0 \mid S = 0\right)\right] \quad (5)$$

The first term is the average treatment effect (ATE), i.e. the rate of return to education for a randomly selected individual. The second term in the square bracket is the OLS *bias* because schooling choice depends on the anticipated potential return, and the bias can be either positive or negative. Therefore, OLS in general doesn't estimate the average treatment effect consistently. From the perspective of individuals who choose college, the OLS bias can be decomposed as follows:

$$E\left(U_1 \mid S = 1\right) - E\left(U_0 \mid S = 0\right)$$
$$= \left[E\left(U_0 \mid S = 1\right) - E\left(U_0 \mid S = 0\right)\right] + E\left(U_1 - U_0 \mid S = 1\right) \quad (6)$$

From the perspective of the unschooled group, the decomposition of the OLS bias is:

$$E\left(U_0 \mid S = 0\right) - E\left(U_1 \mid S = 1\right)$$
$$= \left[E\left(U_1 \mid S = 0\right) - E\left(U_1 \mid S = 1\right)\right] + E\left(U_0 - U_1 \mid S = 0\right) \quad (7)$$

The term in the square bracket in (6) is *selection bias* for college students. It is the mean difference in unobservables between the counterfactual of what a college graduate would earn if he didn't attend college and what an average high school graduate earns. The next term is *sorting gain*, which is the mean gain in the unobservables for college graduates, i.e. the counterfactual difference between what an average college graduate earns and what he would earn if the college degree were not obtained. In Equation (7), the bracketed term is the selection bias for the unschooled group, which is the mean difference in unobservables between the counterfactual of what a high school graduate would earn had he completed college and what an average college graduate earns. The second term is the sorting gain for this group, which is the mean difference in unobservables for high school graduates, i.e. the difference between what an average high

school graduate earns and the counterfactual of what would be earned had he completed college.[10]

Willis and Rosen (1979) show that selection biases can be either positive or negative. When they are both negative, it is consistent with selecting by comparative advantage. On the other hand, positive selection bias in equation (6) and negative selection bias in equation (7) would be consistent with a single-factor (hierarchical) interpretation of ability, i.e. the schooled group on average has higher ability than the unschooled group.[11] It is particularly interesting to note that a positive selection bias for the unschooled group signals possible involuntary selection, meaning that the unschooled group would be better off if they had gone to college, but may be restrained from selecting their preferred alternative by unobserved barriers to college.

Combine the above two types of sorting gains with the average treatment effect, we obtain two parameters that are of great policy interest:

$$
\begin{aligned}
E\left(\beta \mid S=1\right) &= E\left(\ln Y_1 - \ln Y_0 \mid S=1\right) = E\left(\beta\right) + E\left(U_1 - U_0 \mid S=1\right) \\
E\left(\beta \mid S=0\right) &= E\left(\ln Y_1 - \ln Y_0 \mid S=0\right) = E\left(\beta\right) - E\left(U_0 - U_1 \mid S=0\right)
\end{aligned}
\tag{8}
$$

The first equation defines the treatment on the treated effect (TT), and it can be decomposed into the sum of the average treatment effect and the sorting gain for the schooled group. The second equation defines the treatment on the untreated effect (TUT), which is the average treatment effect minus the sorting gain for the unschooled group. The treatment on the treated effect captures the mean gain the schooled group experience, compared with what they would earn if they hadn't gone to college. The treatment on the untreated effect captures the mean gain the unschooled group would experience if they had gone to college, compared with what they earn now. If the sorting gain for the

---

[10] Selection bias compares two groups of persons, the schooled and unschooled, while sorting gain compares two distinct earnings results of the same group. Therefore, the above decompositions by group allow us to extract more information from the data than conventional methods.

[11] For example, if the labor market is dominated by selecting by comparative advantages, then the best lawyers (i.e. schooled or college graduates) are also the worst plumbers (i.e. unschooled or high school graduates), and vice versa. Under the hierarchical ability assumption, however, typical college graduates would be more productive lawyers *and* plumbers than typical high school graduates.

schooled group is positive, it is evidence of purposive sorting based on heterogeneous returns to education.

Selection bias can be obtained from the following alternative decomposition of the OLS estimator:

$$
\begin{aligned}
p\lim\left(\hat{\beta}_{OLS}\right) &= E\left(\ln Y_1 \mid S=1\right) - E\left(\ln Y_0 \mid S=0\right) \\
&= E\left(\beta \mid S=1\right) + \left[E\left(U_0 \mid S=1\right) - E\left(U_0 \mid S=0\right)\right] \\
&= E\left(\beta \mid S=0\right) - \left[E\left(U_1 \mid S=0\right) - E\left(U_1 \mid S=1\right)\right]
\end{aligned}
\tag{9}
$$

Tautologically, the selection bias for the schooled group is the difference between the OLS estimate and treatment on the treated effect, while the selection bias for the unschooled group is the difference between the treatment on the untreated effect and the OLS estimate.

Following Carneiro, Heckman, and Vytlacil (2000), we adopt a two-step procedure to estimate the above parameters. In the first step, a probit model is used to estimate the $\mu_s(Z)$ function of equation (2). The predicted value is called the propensity score, $\hat{P}_i$, where the subscript i denotes each individual. The second step adopts a semi-parametric procedure in which *local polynomial regressions* are used to retrieve the marginal treatment effect. The marginal treatment effect is the marginal gain to schooling of a person just indifferent between taking schooling or not. The marginal treatment effect and parameters derived from it are estimated using the local instrumental variable method developed in Heckman, Ichimura, Todd, and Smith (1998). We do not impose any functional restrictions on the relation between marginal treatment effect and unobservables in the schooling choice equation. Fan (1992, 1993) develops the distribution theory for the local polynomial estimator of $E(\Phi/\Xi=\xi)$, where (Φ,Ξ) is a bivariate random data set. It is shown that $E(\Phi/\Xi=\xi)$ and its first derivative can be consistently estimated by the following algorithm:

$$
\min_{\gamma_1,\gamma_2} \sum_{i \leq N}\left[\Phi_i - \gamma_1 - \gamma_2\left(\Xi_i - \xi\right)\right]^2 G\left(\frac{\Xi_i - \xi}{a_N}\right)
\tag{10}
$$

where N is the sample size. Then, γ_1 is a consistent estimator of $E(\Phi/\Xi=\xi)$, and γ_2 is a consistent estimator of $\partial E(\Phi/\Xi=\xi)/\partial\Xi$. $G(.)$ is a kernel function and a_N is a bandwidth. We use a Gaussian kernel and a bandwidth of 0.2 in the empirical estimation.[12] Intuitively, this algorithm is equivalent to applying weighted least squares at designated point, i.e. $\Xi=\xi$, using all observations but with decaying weights assigned to more distant data points.

More specifically, we estimate a partially linear, conditional expectation model of equation (3)

$$E\left(\beta \mid X, U_s = p\right) = \left(\mu_1\left(X\right) - \mu_0\left(X\right)\right) + E\left(U_1 - U_0 \mid X, U_s = p\right) \qquad (3')$$

By definition the left-hand-side is the marginal treatment effect at $U_s = \mu_s$. We assume linear functional forms for the first term on the right-hand-side of equation (3'), while we estimate the second term, i.e. $E(U_1 - U_0 | X, U_s = p)$ in a nonparametric manner. Following the convention in the literature of semi-parametric estimation (Ichimura and Todd 2004), we first obtain consistent estimates of the linear coefficients with the *double residual regression*, and then retrieve the residuals for the nonparametric estimation. Specifically, we first estimate $E(lnY/P=p)$ and $E(X/P=p)$ with the local polynomial algorithm (i.e. equation (10)). Then we run the *double residual regression* of $lnY - E(lnY/P=p)$ on $X - E(X/P=p)$.[13] This is a simple OLS regression that yields consistent estimates of coefficients of the linear components of equation (1).[14] Let α be the vector of these estimates.[15] Define the nonlinear component residual as $U = lnY - \alpha X$. Use local polynomial regression again to estimate $E(U/P=p)$ and its first derivative. This first derivative by definition is the marginal treatment effect.

The average treatment effect (ATE) is a simple integration (over the support of μ_s) of the MTE with equal weight assigned to each $U_s = \mu_s$. Furthermore, treatment on the

---

[12] This approximates the rule-of-thumb bandwidth selector proposed in Fan and Gilbels (1996).
[13] This procedure is analogous to de-meaning the earnings equations. This approximately purges out the nonlinear components due to the continuity of the nonlinear functions, which allows us to retrieve the nonlinear components later by using the residuals.
[14] We use evenly spaced points from the joint set (with increment equal to 0.01).
[15] Since we pool both the schooled and unschooled groups in this step, we essentially assume the linear components between the two equations in (1) only differ by a constant, which is the average treatment effect. This assumption dramatically simplifies the computation, and it can be easily modified by running separate double residual regression for each group and obtain different α for each group.

treated (TT) and treatment on the untreated (TUT) are simple integration of MTE with the following weighting functions:[16]

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

$$h_{TUT}(u_s) = \frac{\left[\int_{0}^{u_s} f(p)\,dp\right]}{E(1-p)}$$

(11)

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

Intuitively, since we are interested in the marginal individuals whose unobserved heterogeneity of attending college is μ_s, a propensity score that is close to μ_s provides more information than one that is farther away from μ_s. Thus, the observations whose propensity scores are closer to μ_s dominate the estimates. Moreover, since we are interested in the *change* in logarithm of income, i.e. return to schooling, due to an infinitesimal change in μ_s, the first derivative estimator γ_2 in equation (10) consistently estimates the marginal treatment effect.[17]

III. Data and Descriptive Statistics

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).[18] Each wave of the CHIP consists of an urban survey and a

---

[16] For derivations of these weighting functions, see Heckman and Vytlacil (1999, 2000). The TT weight is basically the scaled probability of receiving a propensity score that is greater than μ_s, i.e. being treated. On the other hand, the TUT weight is the scaled probability of receiving a propensity score that is smaller than μ_s, i.e. not being treated.

[17] In practice we also include a quadratic term in equation (10) to improve the accuracy of estimation.

[18] The CHIP-88 and CHIP-95 data are available to the public at the Inter-university Consortium for Political and Social Research (ICPSR). Both data sets have been used intensively by researchers around the

rural survey; we only use the urban survey data for this study. Each urban survey covers thousands of households and individuals in about a dozen provinces.

The three sample years represent distinct phases of economic reform in China. Specifically, 1988 represents the early stage of urban reform that started in 1982 and ended with the 1989 Tian-An-Men Square demonstration. The year 1995 represents the middle stage of urban economic transitions after the reform re-started in 1992 and before the 1997-98 Asian financial crises, and by 2002 economic transition had entered a mature stage. One measure of progress in economic transition is the share of employment in the non-public sector which, as shown in tables 2a, 2b and 2c, increased from 1% in 1988 to 9% in 1995 and then rapidly to 36% in 2002.

Because we are interested in self-selection based on heterogeneous returns, we must assure that individuals in our sample had a reasonable chance of exercising a choice to attend college. The Cultural Revolution (1966-1976) virtually eliminated the choice to attend college. Many youths were sent to the countryside for "rectification" (or "re-education"), and many colleges and even middle schools were either closed or dysfunctional. In 1977, the government reinstated the college entrance exams after a ten-year hiatus. After 1978, all high school graduates who could achieve high enough grades and entrance examination scores could attend college (see Li 2009). As a general rule in the late 1970s, children started elementary school at age 7 and stayed for 5 years; junior high school and senior high school each took 2 years. Thus, an individual who was born in 1962 and started elementary school at age 7 would be a senior in high school in 1978 and could choose to take the required examinations and go to college if he/she performed sufficiently well. Thus we limit our samples to individuals born in 1962 or after.

Our samples are further restricted to working individuals who are living in a household with their parents and who have positive earnings in the survey year. In addition, the sample is limited by the availability of family background information such as parental education and income. As specified above, the samples consist of two

---

world. For details about CHIP-88 and CHIP-95, see Griffin and Zhao (1993) and Riskin, Zhao and Li (2001). CHIP-02 has not yet been released to the public. A recent publication using CHIP-02 is Khan and Riskin (2005).

education groups: 3 or 4-year college graduates and high school graduates.[19] Variable definitions and sample statistics are presented in tables 1, 2a, 2b, and 2c.[20] The proportion of college graduates in the samples was 19% in 1988, 49% in 1995, and 61% in 2002.[21]

We define earnings to include regular wages, bonuses, overtime wages, in-kind wages, and other income from the work unit. The hourly wage rate is calculated as earnings divided by reported hours worked. The nominal average hourly wage almost doubled from 2.30 *yuan* in 1995 to 4.57 *yuan* in 2002 (with negligible inflation), and the increase was larger for college graduates than for high school graduates.[22] The standard deviation of wage rates also doubled. We use parental income five years prior to the survey date as one of the proxies for potential financial constraints on attending college, because this is the closest match we can get to parental income at the time when the individual decided to go to college or not.[23]

IV. Empirical Results

We first examine the probability of acquiring a college education, and then we analyze estimates of various treatment effects.

A. Propensity to Acquire a College Education

Table 3a presents simple probit estimates of college attendance and the mean marginal propensities (probabilities) attributable to each independent variable in the three

---

[19] The education measure includes several degree categories: elementary school or below, junior high, senior high, technical school, junior college (3-year college), and college/university or above. For more details, see Li (2003). Because technical school is different from senior high school and college, we excluded it from our samples. Thus, our samples focus on high school graduates and college graduates.

[20] The sample of 1988 is the largest because we cannot distinguish children and children-in-law in a household. This may cause some problems of mis-matching parents' education and income in the estimation. Yet, this problem should not be very serious for 1988 because the oldest age should be 26 years old, still somewhat too young to be married. In CHIP-95 and CHIP-02, the data can distinguish children and children-in-law.

[21] Note these proportions are calculated conditional on high school graduates and with the urban samples. Therefore they are significantly higher than those calculated using nationwide population survey.

[22] The nominal exchange rate between the U.S. dollar and the Chinese *yuan* depreciated sharply during our sample period. It was 3.7 *yuan*/dollar in 1988, 8.4 in 1995, and 8.3 in 2002.

[23] Such information is not available in CHIP-88, so current parental income is used instead.

sample years, 1988, 1995, and 2002, respectively. The regressors include those related to the budget constraint. In particular, parental income provides the financial resources to attend college.[24] Since individuals in the sample are currently employed, the time they chose to enter college is at least four years prior to the date of the survey. In CHIP-95 and CHIP-02, we have information on parental income up to five years prior to the survey date, and we use parental income five years before the survey as a proxy for any financial constraint affecting college attainment. We also include the number of siblings in the household as a proxy for a financial constraint, as children are likely to compete for financial resources to fund education.[25]

Parental education is included as it may be related to ability formation, attitude toward college, and information possessed about going to college, because (i) parental educational achievement is likely to reflect parental ability, which may be inherited by offspring; (ii) parental education is likely to contribute to child-rearing practices that foster pre-college "ability" attainment and to foster a generally positive attitude toward higher education (Carneiro, Meghir, and Parey, 2007). We cannot rule out, though, that the financial ability to provide childhood investments in learning and nutrition may affect measured ability at older ages (Heckman and Li, 2004).

We include a birth year dummy to capture year-specific factors related to opportunities of going to college, such as admission quotas administered by the Chinese Government Ministry of Education. This has been an effective cap on admissions, and from 1977 to 1999 the admission rate rose from less than 10% to 48% (Li, 2009). Other controls include dummy variables for gender dummy and for the ethnic minority.

In all three years, both parental income and education exert a positive impact on children's chances of attending college, and in most cases the estimates are statistically significant. This result implies that parental income and education play distinct roles in children's education attainment despite their high correlation. In all three years, father's

---

[24] College students generally do not have jobs in China.
[25] Unfortunately, it is an imperfect measure of household size, as not all children lived in the household during the time of survey.

education has a larger effect than that of mother's. The largest difference is found in 1988, but the difference becomes much smaller and negligible in 1995 and 2002.[26]

Mother's income shows a much larger effect on college choice than that of father's in 1988 and 1995, but not in 2002. The marginal effect of mother's income is about four times larger than that of father's in 1988 and two times larger in 1995. The convergence of the marginal effects of father's and mother's income is consistent with parental income's having more strongly reflected ability than financial constraints in the earlier period, when tuition was not charged for most colleges and family financial resource wasn't a barrier to college attendance. We postulate that, given father's income, higher mother's income supported better nutrition that contributed to ability formation in children. However, when ability to pay became a significant barrier to college attendance and nutritional problems due to inadequate diet diminished with general economic growth, the impact of mother's income approached that of the father's at the margin.

We can use the estimated marginal effects to evaluate the relative impacts of the variables that influence schooling choice. For example, in 1995, the marginal impact of an additional year of father's education has the same impact as an increase of 5.7 thousand *yuan* of father's income. However, an additional year of father's education is only "worth" 2.6 thousand *yuan* in 2002, implying a rise in the importance of parental income relative to parental education. We attribute this change to tuition hikes during the 1990s. The pattern is not as pronounced for mother's income, though.

The estimated effect of another proxy for family financial constraint — the number of children in the household — is large, negative, and statistically significant in 1988. One more sibling reduces the probability of attending college by 4.5%. In 1995 the impact is still a negative 3.7% (and almost statistically significant the 10% level), while in 2002 the impact becomes insignificant. Although this decline in the marginal impact of an additional child would appear to contradict our hypothesis that the influence of financial constraints on college attendance increased over time, we believe that it is due to increasingly stringent enforcement of the one-child policy which substantially reduced variation in the number of children among urban households. Ethnic minority status does

---

[26] The results for 1988 should be interpreted with caution because of possible mismatch of parental education and income.

not appear to be very important in college choice either, although it became quite negative and almost significant at the 10% level in 2002.

Although it may at first appear surprising that in 1995 and 2002, the coefficient on the gender (male) dummy is negative and statistically significant at the 10% level, we interpret this higher likelihood for females to attend college as the result of selectivity prior to high-school attendance. In all three samples, female students comprise a smaller proportion of high-school graduates than male students. For female students, enrolling in high school signals strong commitment to attempting college; female students who have completed senior high school are more likely to continue into college.

In table 3a we compare marginal coefficients across years using sample means for each year. In table 3b we perform the same exercise using the overall sample means, i.e. the three-year average. In order to anchor the birth year dummy, we choose the cohort born in 1968 which appears in all three samples,[27] and we deflate nominal parental income by the urban CPI. Table 3b shows that, for one thousand *yuan* increase in the real value of father's income, the probability of going to college increases by 0.5 percentage point in1988 and 4 percentage points in 2002. In 2002, the effect of father's income surpasses education and becomes the dominant factor in college entrance. The impact of mother's income also increases, but by a smaller amount. The growing importance of father's income for college entrance, controlling for parental education, is consistent with the rising impact of higher college costs. The marginal effect of parental education increased sharply from 1988 to 1995, and then declined moderately in 2002. An additional year of father's education increased the probability of going to college by 1 percentage point in 1988, 3.4 percentage points in 1995, and 2.4 percentage points in 2002. The effect of mother's education displays the same pattern but the drop between 1995 and 2002 is smaller.

The probit models generate a propensity score for each observation, which is the predicted probability of college attendance. The frequency distributions of these propensities provide a reduced-form picture of growing college attendance in China

---

[27] For 1988, those born in 1968 or before were combined into the same cohort.

(Figure 2).[28] 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 graduates and high school graduates. 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 at least a college education 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 high school graduates is supported over a range of propensity scores from approximately zero through nearly 0.6;[29] 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 college graduates is supported over the range of propensities between approximately zero and 0.7 in 1988, approximately zero and greater than 0.9 in 1995, and from about 0.1 through 1.0 in 2002.

Examining these distributions more carefully reveals some interesting trends. Table 4 shows that in 1988, 19.2% of the sample were college graduates and had a propensity score equal to or greater than 0.323. We define the *cut-off propensity* as the propensity score corresponding to the cumulative frequency of the total sample that had graduated from college in the sample year. If there were no unobserved heterogeneity in returns to schooling, and if there were no financial constraints on college attendance, then the frequency distributions of propensity scores of college attenders and non-attenders would not overlap. In 1988 10.5% of the entire sample had scores higher than the cutoff score and yet did not go to college; by construction the same fraction of the entire sample had scores lower than this value, yet they did attend college. This ratio becomes 15.9% in 1995 and 15.3% in 2002. Moreover, the percentage of such "misfits" in the unschooled group has increased from 12.9% in 1988, to 27.2% in 1995 and 39.7% in 2002, while in the schooled group this percentage has decreased from 54.6% in 1988, to 38.3% in 1988 and 24.9% in 2002.

---

[28] The sample densities are smoothed with Gaussian kernels with optimal bandwidths defined in Silverman (1986).
[29] A small support implies that omitted factors play important roles or large unobservable heterogeneity exists in schooling decision.

These patterns suggest that unobserved heterogeneity increased dramatically over our sample period, mostly between 1988 and 1995, and such increased heterogeneity is apparently affecting the unschooled group more than the schooled group. The increased heterogeneity could reflect (1) a growing proportion of agents with high propensity scores who cannot realize their high potential because they are unable to finance college education; or (2) a growing importance of unobserved comparative advantage. If (1) dominates, then we should observe selection bias and sorting gain diminishing over time for the schooled, and increasing over time for the unschooled; however, if (2) dominates, sorting gains for both groups should increase.

B. College Education and Earnings

Table 5 contains the results of OLS, IV, and semi-parametric local instrumental estimation of the effect of college attendance on earnings. For each sample we present two specifications of the earnings equation. The benchmark specification includes both parental income and education, and the comparison specification only uses parental income. The OLS estimates are similar to those reported elsewhere for comparable time periods, and exhibit an upward trend in private returns to college education.[30] The IV estimates of the return to college education (all of which use the propensity score as the instrument for college attendance) are smaller than the OLS estimates in 1988 but become considerably higher than the OLS estimates in 2002. Since in general neither OLS nor the conventional IV method consistently estimates the average treatment effect, such differences between OLS and IV do not have clear implications (Caneiro, Heckman, and Vytlacil, 2000).[31] In fact, both OLS and IV underestimate the ATE in the benchmark specification in all three samples, and the OLS bias increased from a negligible -1.5% in 1988, to -13.5% in 1995 and -72.9% in 2002 (significant at 1%).

We now turn to our estimates of returns to schooling based on the semi-parametric local IV estimation. Variables included in the choice equation but excluded

---

[30] See Fleisher and Wang (2004) and Li (2003) for estimates and a summary of other studies for the same period.

[31] In the literature of homogenous return, IV estimates are usually found to be higher than the OLS estimates after correcting for omitted ability bias. The explanation is attenuation bias caused by measurement errors (Li and Luo 2004, Butcher and Case 1994, and Ashenfelter and Zimmerman 1997).

from the earnings equation *de facto* serve as instruments. In the benchmark model the number of siblings and birth year are instruments for ability, while in the comparison model parental education is an additional instrument. Such choice of instruments has a long pedigree in the literature (e.g. Mare 1980) and is similar to that of Heckman, Urzua, and Vytlacil (2006) and Heckman and Li (2004).[32] The results from both specifications are generally consistent and robust.

Based on the semi-parametric estimation, we find that between 1988 and 2002 the average treatment effect — the return to education for a randomly selected individual — has increased substantially. In the benchmark specification, the cumulative rate of return for four years of college has increased from an insignificant 24.4% in 1988,[33] to a insignificant 42.0% in 1995, and then to a very significant 165.1% in 2002.[34] However, when this dramatic change is decomposed into treatment on the treated (TT) and treatment on the untreated (TUT), we obtain strikingly contrasting results. The component of returns, TT, the *realized* return that is achieved by individuals who actually completed four years of college, changed from 105.6% in 1988, to 48.6% in 1995, and to 175.1% in 2002.[35] This return compares the actual earnings of college graduates with the counterfactual earnings that they would have received as high-school graduates. TUT, representing the *unrealized* counterfactual return that could have been earned as college graduates by those who did not go to college, changed from 10.4% to 37.4%, and then to 149.7%. The increase in the unrealized return outpaces that of the realized return. This implies that although the return to those who go to college has increased drastically since 1988, the potential return for those who do not (or cannot) go to college has increased even more.

We obtain further insights by analyzing the changes in selection bias and sorting gain for the schooled and unschooled groups over time. For those who go to college, the

---

[32] Heckman, Urzua, and Vytlacil (2006) uses number of siblings and mother's graduation status as instruments; while Heckman and Li (2004) uses birth year and parental education as instruments.

[33] 24.4% is computed with the formula 100[exp(0.218)-1] where 0.218 is taken directly from the corresponding entry in Table 5. All estimates discussed in this paragraph are transformed this way.

[34] These are four-year accumulative rates. The corresponding annualized rates are 5.6%, 9.2%, and 27.6% respectively.

[35] In a few cases, the result for 1995 does not follow the trend between 1988 and 2002. One reason is that for 1988, the wage measure is based on annual labor income; while for 1995 and 2002, the wage measure is based on hourly wage. Thus, the results for 1995 and 2002 are generally more comparable.

selection bias, the mean difference in unobservables between the counterfactual of what a college graduate would earn if he didn't attend college and what an average high school graduate earns, became more negative (-0.517, -0.180, and -0.766, respectively). That is, those who go to college are increasingly drawn from a pool who would have been below-average earners among high-school graduates. Negative selection bias is evidence of self-selection based on comparative advantage, and the negative trend in this variable implies that the degree of this self-section among college graduates has increased. The sorting gain for college graduates, the counterfactual difference between what an average college graduate earns and what he/she would earn if had not gone to the college, is small and has diminished from 0.502 in 1988 to 0.037 in 2002.

For those who do not go to college, the selection bias, the mean difference in unobservables between the counterfactual of what a high school graduate would earn had he completed college and what an average college graduate earns has changed from negative to positive and become larger in absolute value (-0.104, 0.102, and 0.669, respectively). This trend in self-section bias for high-school graduates contradicts the evidence of increasing self-selection inferred from the trend in selection bias for college graduates. The negative selection bias for college graduates and positive selection bias for high-school graduates contradicts the hypothesis of single-factor (hierarchical) ability. College graduates do not appear to be better in everything than those who stop their schooling with high school graduation. Rather, it appears that in later years some high school graduates could not self-select into college despite the high potential return. The sorting gain for the unschooled group has been small and hardly significant (0.119, 0.033, and 0.060, respectively), similar to the pattern of sorting gain for college graduates.[36]

The above findings are consistent with the following unified explanation. Over the sample period, most college graduates are still college-worthy, but a growing number of college-worthy high school graduates have been denied the opportunity to attend

---

[36] Our small estimates of sorting gain are not inconsistent with our estimates of selection bias. Negative selection bias indicates that the agent's "ability" known to the agent but not known by the econometrician is below average in the schooling class not chosen. But it does NOT imply that the agent receives above-average earnings in the chosen schooling class. Positive selection bias indicates that the agent is (would be) "above-average" among the members of schooling group not chosen. In general the sign of sorting gain cannot be predicted from the sign of selection bias.

college. This explanation is consistent with the decline in sorting gains for both the schooled and unschooled groups. We could say that the group of high school graduates has become increasingly "contaminated" by college-worthy students. Aggregate data support this explanation. Throughout our sample period, the college admission rate increased from only 3% in 1977 to nearly 40% in 2000, largely because of the expansion of college enrolment capacity. However, the tuition hikes that started in 1989 have caused financial difficulty for many families in funding college education. From 1990 to 1998, the ratio of average tuition per student to per capita GDP increased from 1.54% to 34% (Li, 2009).

The heterogeneous return model postulates that those who attend college do so because they benefit more than those who choose not to attend, and thus the college choice is an endogenous response to such a benefit. It is consistent with someone choosing not to attend college because financial or psychic costs are expected to outweigh financial gains (Carneiro, Heckman, and Vytlacil 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.* Figure 3 depicts the estimated MTE of college education from the benchmark specification of the earnings equation for the years 1988, 1995, and 2002. The MTE captures the observed gross financial gains from attending college, and it is in this sense that we identify those with high MTE as most "college-worthy". The MTE curves for 1988 support the hypothesis that people with high gross financial returns are also more likely to attend college and those with smaller expected financial returns are less likely to attend college.

However, the MTE curves become U-shaped in 1995 and 2002. This is consistent with the explanation given above, that only some college-worthy students are likely to attend college (as shown towards the left portion of the curve) while other college-worthy students are less likely to attend college (as shown towards the right portion of the curve). A U-shaped MTE is inconsistent with the joint hypothesis that agents' unobserved heterogeneity involves only their comparative advantage to benefit from more schooling.

It is consistent with an unobserved barrier to college attendance in China, e.g. psychic costs or unobserved financial barriers (Carneiro, Heckman, and Vytlacil 2004, p. 25).[37]

V. Conclusion

In this study we investigate the returns to education during China's economic transition, using a semi-parametric technique that is based on the assumptions of heterogeneous returns and that individuals act on the anticipated return. We estimated policy relevant parameters including the treatment effect on the treated, treatment effect on the untreated, and the average treatment effect, and discussed their dynamics for the period from 1988 to 2002.

Although all three estimation methods — OLS, IV, and semi-parametric local IV (SPIV) — reveal a substantial increase in returns to schooling in China between 1988 and 2002, they differ substantially in the estimated levels of returns. We believe that only the SPIV estimates answer well-posed policy questions. Our results based on SPIV are quite robust for different specifications.

We find that the increase in average treatment effect is driven by both an increase in the returns to those who have chosen college education, but also by a larger increase in the potential returns to those who remain "unschooled". Our estimates on selection bias indicate that while purposive selection has increased for college graduates, it has declined for high school graduates. In addition, sorting gain has become less pronounced for college graduates and remained small for high school graduates. We interpret these results as evidence indicating that while the higher education in China has increased reward for the schooled group, perhaps reflecting the rising influence of market forces, it has become less efficient in terms of financing college education among some college-worthy youth.

Our sample covers a period in which China's higher education system underwent major structural changes. Higher costs of college education affect self-selection in two

---

[37] When we add parental education as additional instrument, the results are quite robust. Specifically, the changing pattern of MTE over time is strikingly similar. These results are not presented here but available upon request.

ways. Individuals (and their families) in the schooled group have responded to higher expected returns and have willingly paid the higher costs of choosing college. On the other hand, among the unschooled group, we find evidence that individuals who would reap a return more than sufficient to compensate for the costs of college attendance have chosen not to go to college.

This finding suggests that either the distaste for college education has increased over time, or that financial constraints on college attendance have become more severe. We believe that the second explanation is likely in light of the changes in education finance in China. Thus, it seems to us that the movement toward higher tuition and private funding of higher education, while justified on many grounds, will also contribute to increasing income inequality in a vicious cycle. More specifically, those from wealthy families are more likely to reap the higher returns of education and then will become wealthier; however, those from poor families may be excluded from schooling opportunities and thus remain poor. Therefore, government policies that help individuals from financially-disadvantaged families gain access to higher education become crucial for equal opportunity to the benefits of reforms in China.

Labor market reform is a critical component of the transition from planning to markets, and the old wage-grid and *nomenklatura* systems are rapidly being replaced by a market system that reflects the true value of education and allows individuals to make the schooling choices that they deem to be optimal. Since the economic reform started, two major changes in the China's higher education system have been enrollment expansion and tuition hike. Based on our results, we find that increased opportunities for college education raised the benefits for those who attended college. Yet, the increasing financial burden of college tuition hindered higher education for some college-worthy youth, and the lost gain associated with not attending college for this group has increased significantly as well. The main policy implication of our results is that labor market and education system reform without accompanying capital market reform deprives the financially disadvantaged of the potential economic benefits.

References

Ashenfelter, Orley, and David J. Zimmerman, "Estimates of Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers," *Review of Economics and Statistics* 79: 1 (1997), 1-9.

Björklund, Anders, and Robert Moffitt, "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *Review of Economics and Statistics* 69:1 (1987), 42-49.

Brainerd, Elizabeth, "Winners and Losers in Russia's Economic Transition," *American Economics Review* 88:5 (December 1998), 1094-1116.

Butcher, Kristin F. and Anne Case (1994), "The Effects of Sibling Composition on Women's Education and Earnings", *Quarterly Journal of Economics* 109, 1994, 443-450.

Card, David, "The Causal Effect of Education on Earnings", in: Ashenfelter, Orler and Card, David, eds., *Handbook of Labor Economics*, North Holland-Elsevier Science Publisher, volume 3A, 1999, Chapter 30:1801-1863.

Carneiro, Pedro, and James J. Heckman, "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal* 112 (October 2002), 705-734.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil, "Estimating the Returns to Education when It Varies Among Individuals," University of Chicago working paper (2000).

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil, "Understanding What Instrumental Variables Estimate: Estimating the Average and Marginal Returns to Schooling," University of Chicago working paper (2003).

Carneiro, Pedro, Costas Meghir, and Matthias Parev, "Maternal Education, Home Environments and the Development of Children and Adolescents," IZA Discussion Papers 3072 (2007).

Fan, J., "Design-adaptive Nonparametric Regression," *Journal of the American Statistical Association* 87 (1992), 998-1004.

Fan, J., "Local Linear Regression Smoothers and Their Minimax Efficiencies," *The Annals of Statistics* 21 (1993), 196-216.

Fan, J., and I. Gijbels, *Local Polynomial Modeling and Its Applications* (Chapman & Hall, 1996).

Fleisher, Belton, "Higher Education in China: A Growth Paradox," in Yum K. Kwan, and Eden S. H. Yu (Eds.), *Critical Issues in China's Growth and Development* (Ashgate Publishing Ltd., Aldershot, UK, 2005), 3-21.

Fleisher, Belton, and Jian Chen, "The Coast-Noncoast Income Gap, Productivity, and Regional Economic Policy in China," *Journal of Comparative Economics* 25:2 (1997), 220-36.

Fleisher, Belton, Keyong Dong, and Yunhua Liu, "Education, Enterprise Organization, and Productivity in the Chinese Paper Industry," *Economic Development and Cultural Change* 44:3 (April 1996), 571-587.

Fleisher, Belton, Klara Sabirianova, and Xiaojun Wang, "Returns to Skills and the Speed of Reforms: Evidence from Central and Eastern Europe, China, and Russia," *Journal of Comparative Economics* 33:2 (June 2005), 351-370.

Fleisher, Belton, and Xiaojun Wang, "Skill Differentials, Return to Schooling, and Market Segmentation in a Transition Economy: The Case of Mainland China," *Journal of Development Economics* 73:1 (2004), 715-728.

Fleisher, Belton, and Xiaojun Wang, "Returns to Schooling in China under Planning and Reform," *Journal of Comparative Economics* 33:2 (2005), 265-277.

Fleisher, Belton, and Dennis Tao Yang, "China's Labor Markets," Ohio State University working paper (2003).

Giles, John, Albert Park, and Juwei Zhang, "The Great Proletarian Cultural Revolution, Disruption to Education, and Returns to Schooling," Michigan State University working paper (2004).

Griffin, K., and R. Zhao (Eds.), *The Distribution of Income in China* (New York, St. Martin's Press, 1993).

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45:1 (January 1977), 1-22.

Hannum, Emily, and Meiyan Wang, "Geography and Educational Inequality in China," *China Economic Review* 17: 3 (2006), 253-265.

Heckman, James J., "China's Human Capital Investment," *China Economic Review* 16:1 (2005), 50-70.

Heckman, James, Hidehiko Ichimura, Petra Todd, and Jeffrey Smith, "Characterizing Selection Bias Using Experimental Data," *Econometrica* 66:5 (1998), 1017-1098.

Heckman, James J., and Xuesong Li, "Selection Bias, Comparative Advantage, and Heterogeneous Returns to Education: Evidence from China in 2000," *Pacific Economic Review* 9 (2004), 155-171.

Heckman, James J., Sergio Urzua, and Edward Vytlacil, "Understanding Instrument Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics* 88, 3 (August 2006): 389-432.

Heckman, James J., and Edward Vytlacil, "Local Instrumental Variable and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences* 96 (1999), 4730-4734.

Heckman, James J., and Edward Vytlacil, "Local Instrumental Variables," in C. Hsian, K. Morimune, and J. Powells (Eds.), *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Eco0nomic Theory and Econometrics: Essays in Honor of Takeshi Amemiya* (Cambridge: Cambridge Unviersity Press, 2000), 1-46.

Ichimura, Hidehiko, and Petra E. Todd, "Implementing Nonparametric and Semiparametric Estimators," manuscript prepared for *Handbook of Econometrics* 5, 2004.

Jones, Derek C., and K. Ilayperuma, "Wage Determination under Plan and Early Transition: Bulgarian Evidence using Matched Employer-Employee Data," *Journal of Comparative Economics* 33:2 (June 2005), 227-243.

Khan, Azizur R., and Carl Riskin, "China's Household Income and Its Distribution, 1995 and 2002," *The China Quarterly* 182 (2005), 356-384.

Knight, John, and Lina Song, "The Determinants of Urban Income Inequality in China," *Oxford Bulletin of Economics and Statistics* 53:2 (1991), 123-154.

Li, Haizheng, "Economic Transition and Returns to Education in China," *Economics of Education Review* 2 (2003), 317-328.

Li, Haizheng and Yi Luo, "Reporting Errors, Ability Heterogeneity, and Returns to Schooling in China?" *Pacific Economic Review*, Special Issue edited by Orley C. Ashenfelter and Junsen Zhang, Vol. 9, pp.191-207, 2004.

Li, Haizheng, "Higher Education in China: Complement or Competition to American Universities," NBER working paper, 2009, http://www.nber.org/~confer/2008/augm08/li.pdf

Liu, Zhiqiang, "Earnings, Education, and Economic Reform in China," *Economic Development and Cultural Change* 46:4 (1998), 697-726.

Mare, R. D., "Social Background and School Continuation Decisions," *Journal of the American Statistical Association* 75: 370 (1980), 295-305.

Meng, Xin, *Labour Market Reform in China* (Cambridge, UK: Cambridge University Press, 2000).

Meng, Xin, and R. G. Gregory, "The Impact of Interrupted Education on Subsequent Educational Attainment: A Cost of the Chinese Cultural Revolution," *Economic Development and Cultural Change* 50:4 (2002), 934-955.

Meng, Xing, and Michael P. Kidd, "Labor Market Reform and the Changing Structure of Wage Determination in China's State Sector during the 1980s," *Journal of Comparative Economics* 25:3 (1997), 403-421.

Munich, Daniel, Jan Svejnar, and Katherine Terrell, "Returns to Human Capital under the Communist Wage Grid and During the Transition to a Market Economy," *Review of Economics and Statistics* 87:1 (2000), 100-123.

Orazem, Peter, and Milan Vodopivec, "Winners and losers in the Transition: Returns to Education, Experience, and Gender in Slovenia," *The World Bank Economic Review* 9:2 (1995), 201-230.

Riskin, C., R. Zhao, and S. Li (Eds.), *China's Retreat from Equality: Income Distribution and Economic Transition* (Armonk, N.Y., M.E. Sharpe, 2001).

Roy, A., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (1951), 135-146.

Silverman, B. W., *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, 1986).

Wang, Sangui, Dwayne Benjamin, Loren Brandt, John Giles, Yingxing Li, and Yun Li, "Inequality and Poverty in China during Reform," PMMA working paper (2007).

Willis, R., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics* (Amsterdam: North-Holland, 1986).

Willis, R., and S. Rosen, "Education and Self-Selection," *Journal of Political Economy* 87:5 (part 2, 1979), S7-36.

Yang, Dennis Tao, "Urban-Biased Policies and Rising Income Inequality in China," *American Economic Review* 89:2 (1999), 306-310.

Yang, Dennis Tao, "What Has Caused Regional Inequality in China?" *China Economic Review* 13:4 (2002), 331-334.

Yang, Dennis Tao, "Determinants of Schooling Returns during Transition: Evidence from Chinese Cities," *Journal of Comparative Economics* 33:2 (2005), 244-264.

Zhang, Junsen, and Yaohui Zhao, "Economic Returns to Schooling in Urban China, 1988-1999," paper presented at the meetings of the Allied Social Sciences Association, Washington DC, 2002.

Zhou, Xueguang, "Economic Transformation and Income Inequality in Urban China: Evidence from Panel Data," *American Journal of Sociology* 105:4 (January 2000), 1135-74.

Zhou, Xueguang, and Liren Hou, "Children of the Cultural Revolution: The State and the Life Course in the People's Republic of China," *American Sociological Review* 65:1 (February 1999), 12-36.

Zhou, Xueguang, and Phyllis Moen, *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]. 2002.

Tables and Figures

Table 1.—Variable Definition

Variable	1988	1995	2002
Father Education (years)		Father's Education	
Mother Education (years)		Mother's Education	
Father's Salary in Earlier Years (1000 Yuan)	Father's Annual Salary in 1988	Father's Annual Salary in 1990	Father's Annual Salary in 1998
Mother's Salary in Earlier Years (1000 Yuan)	Mother's Annual Salary in 1988	Mother's Annual Salary in 1990	Father's Annual Salary in 1998
Number of Children		Number of Children living in the household	
Gender (male=1)		Gender dummy	
Ethnic		1 if the individual is an ethnic minority (non-Han Chinese)	
Work Experience (years)	Estimated by age minus years of schooling minus 6	Year of Work Experience Reported	
Wage	Monthly Wage	Hourly Wage Rate	
Government Sector	Government or public institutions	Not available	
State-owned Sector		State-owned at central or provincial governmental level	
Local Publicly-owned Sector		Publicly-owned at lower government level	
Urban Collective Sector		Collectively Owned Sector	
Province		Dummy variables for each province	
Industry		Dummy variables for each industry	
Birth Year		Dummy variables for the year of birth	
College		1 if individual is a college graduate	

Note: wage includes regular wage, bonus, subsidies and other income from the work unit.

Table 2a.—Descriptive Statistics for 1988 Sample

Variable	Full Sample		College Graduates		High School Graduates	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Father Education (years)	10.19	3.40	11.71	3.50	9.83	3.28
Mother Education (years)	8.33	3.45	9.39	3.83	8.08	3.31
Father Salary 1988 (1000 Yuan)	2.12	1.31	2.30	1.56	2.08	1.24
Mother Salary 1988 (1000 Yuan)	1.22	1.06	1.39	1.08	1.19	1.06
Number of Children	1.83	0.85	1.64	0.73	1.88	0.87
Gender (Male=1)	0.50	0.50	0.51	0.50	0.50	0.50
Ethnic	0.03	0.18	0.04	0.19	0.03	0.18
Monthly Salary (1000 Yuan)	0.15	0.48	0.26	0.44	0.13	0.49
Work Experience (years)	4.26	2.35	2.71	1.68	4.63	2.34
State-owned Sector	0.42	0.49	0.52	0.50	0.40	0.49
Local Publicly-owned Sector	0.40	0.49	0.42	0.49	0.40	0.49
Urban Collective Sector	0.17	0.37	0.06	0.24	0.19	0.39
College	0.19	0.39	-	-	-	-
Number of Observations	1128		216		912	

Note: the omitted ownership sector is non-public sector including private enterprises, Sino-foreign joint venture, etc.

Table 2b.—Descriptive Statistics for 1995 Sample

Variable	Full Sample		College Graduates		High School Graduates	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Father Education (years)	11.70	3.27	12.89	2.97	10.86	3.21
Mother Education (years)	9.80	3.50	11.04	3.35	8.91	3.33
Father Salary 1990 (1000 Yuan)	3.55	2.13	3.76	2.49	3.41	1.83
Mother Salary 1990 (1000 Yuan)	2.58	1.55	2.87	1.77	2.38	1.34
Number of Children	1.66	0.65	1.61	0.64	1.70	0.65
Gender (male=1)	0.60	0.49	0.59	0.49	0.62	0.49
Ethnic	0.04	0.19	0.04	0.19	0.03	0.18
Hourly Wage (Yuan/hour)	2.30	1.84	2.69	2.21	2.02	1.46
Work Experience (years)	5.55	3.77	5.02	3.55	5.94	3.87
State-owned Sector	0.26	0.44	0.32	0.47	0.21	0.41
Local Publicly-owned Sector	0.55	0.50	0.52	0.50	0.56	0.50
Urban Collective Sector	0.10	0.31	0.06	0.24	0.13	0.34
College	0.42	0.49	-	-	-	-
Number of Observations	686		285		401	

Note: the omitted ownership sector is non-public sector including private enterprises, Sino-foreign joint venture, etc.

Table 2c.—Descriptive Statistics for 2002 Sample

Variable	Full Sample		College Graduates		High School Graduates	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Father Education (years)	10.57	3.22	11.23	3.25	9.53	2.89
Mother Education (years)	9.54	3.06	10.04	2.94	8.75	3.08
Father Salary 1998 (1000 Yuan)	10.24	6.54	11.17	7.27	8.75	4.82
Mother Salary 1998 (1000 Yuan)	6.97	4.74	7.52	5.29	6.10	3.54
Number of Children	1.26	0.46	1.27	0.47	1.24	0.45
Gender (male=1)	0.61	0.49	0.54	0.50	0.71	0.45
Ethnic	0.05	0.22	0.04	0.20	0.07	0.25
Work Experience (years)	6.46	4.96	5.82	4.54	7.46	5.42
Hourly Wage (Yuan/hour)	4.57	3.64	5.21	4.16	3.56	2.28
Government Sector	0.30	0.46	0.42	0.49	0.11	0.31
State-owned Sector	0.12	0.33	0.10	0.30	0.15	0.36
Local Publicly-owned Sector	0.16	0.37	0.13	0.34	0.22	0.41
Urban Collective Sector	0.06	0.24	0.03	0.18	0.11	0.31
College	0.61	0.49	-	-	-	-
Number of Observations	654		402		252	

Note: the omitted ownership sector is non-public sector including private enterprises, Sino-foreign joint venture, share-holding company, etc.

Table 3a.—Propensity Estimates

Variable	1988 (based on current year parental income)				1995 (based on parental income in 1990)				2002 (based on parental income in 1998)			
	Param.	t-ratio	p-value	Mean Marginal Effect	Param.	t-ratio	p-value	Mean Marginal Effect	Param.	t-ratio	p-value	Mean Marginal Effect
Constant	-2.475	-10.043	0.000		-1. 804	-6.182	0.000		-1.847	-4.588	0.000	
Father Education	0.082	4.849	0.000	**0.020**	0.087	4.543	0.000	**0.034**	0.063	3.288	0.001	**0.023**
Mother Education	0.002	0.132	0.448	**0.001**	0.074	4.030	0.000	**0.029**	0.057	2.784	0.003	**0.021**
Father Salary	0.028	0.794	0.214	**0.007**	0.015	0.533	0.297	**0.006**	0.026	2.230	0.013	**0.010**
Mother Salary	0.098	1.989	0.023	**0.024**	0.032	0.776	0.219	**0.012**	0.021	1.400	0.081	**0.008**
Number of Children	-0.222	-3.664	0.000	**-0.045**	-0.098	-1.201	0.115	**-0.037**	0.023	0.191	0.424	**0.009**
Gender	-0.087	-0.922	0.178	**-0.020**	-0.140	-1.297	0.098	**-0.054**	-0.539	-4.630	0.000	**-0.195**
Ethnic	0.072	0.287	0.387	**0.017**	0.177	0.643	0.260	**0.070**	-0.295	-1.221	0.111	**-0.114**
Number of Observations	1128				686				654			
Log Like.	-466.300				-407.807				-366.503			
Like. Ratio	0.154				0.124				0.159			
Pseudo-R2	0.316				0.510				0.690			

Notes: The dependent variable is binary, which is 1 for graduated from 3- or 4-year college and 0 otherwise. For 1995 and 2002, father's and mother's salary is for the time 5 years prior to the sample year; while for 1988, they are based on the current year income due to data limitation. The results for birth year dummies are not reported. The marginal effects are calculated using (mean+1) for easy interpretation, based on the sample average of each year. For dummy variables, the marginal effects are calculated based on changing its value from 0 to 1.

Table 3b.—Marginal Effect Estimates

Variable	1988 Mean Marginal Effect	1995 Mean Marginal Effect	2002 Mean Marginal Effect
Father Education	0.010	0.034	0.024
Mother Education	0.0002	0.029	0.021
Father Salary	0.0047	0.022	0.040
Mother Salary	0.020	0.043	0.032
Number of Children	-0.022	-0.037	0.009
Gender	-0.010	-0.054	-0.20
Ethnic	0.009	0.069	-0.12

Notes: The marginal effects are evaluated at the overall
sample mean (three years) for each variable. It is calculated
as (mean+1) in order to have a more meaningful
interpretation. For dummy variables, the marginal effect is
calculated based on the difference between 0 and 1. For birth
year dummies, we use the 1968 cohort as the default. The
marginal effects for father and mother's salary are measured
by 1,000 Yuan increase in 1984 value based on the urban CPI
index (1988, 157.2; 1995, 358.4; 2002, 396.3).

Table 4.—Comparison of Propensity Distributions

|	1988	1995	2002
Sample size	1128	686	654
Number of non-attenders	912	401	252
Number of attenders	216	285	402
Proportion of sample who are college attenders or graduates	19.15%	41.55%	61.47%
Cut-off Propensity	0.323	0.459	0.554
Number of respondents in the "wrong" group	118	109	100
Percentage of the non-attender group	12.94%	27.18%	39.68%
Percentage of the attender group	54.63%	38.25%	24.88%
Percentage of the total sample	10.46%	15.89%	15.29%

Notes: The cut-off propensity is the propensity score that corresponds to the cumulative frequency of the total sample that was attending or had graduated from college in the sample year.

Table 5: Marginal Treatment Effect Estimates

| Parameter	CHIP88		CHIP95		CHIP02	
Instrument Set	Benchmark	+ Educ.	Benchmark	+ Educ.	Benchmark	+ Educ.
OLS	0.203[a] (0.030)	0.195[a] (0.042)	0.216[a] (0.056)	0.234[a] (0.051)	0.246[a] (0.053)	0.257[a] (0.054)
IV*	0.135[a] (0.053)	0.078 (0.064)	0.052 (0.299)	0.353[b] (0.164)	0.632[a] (0.182)	0.603[a] (0.169)
ATE	0.218 (0.207)	0.061 (0.169)	0.351 (0.391)	0.523[b] (0.266)	0.975[a] (0.305)	0.836[a] (0.270)
TT	0.721[a] (0.200)	0.417[b] (0.240)	0.396 (0.483)	0.610[c] (0.371)	1.012[a] (0.312)	0.896[a] (0.325)
TUT	0.099 (0.246)	-0.024 (0.202)	0.318 (0.426)	0.461[c] (0.326)	0.915[b] (0.432)	0.740[b] (0.321)
Bias =OLS-ATE $=E(U_1|S=1)-E(U_0|S=0)$	-0.015 (0.196)	0.134 (0.159)	-0.135 (0.394)	-0.289 (0.258)	-0.729[a] (0.307)	-0.579[b] (0.259)
Selection Bias (S=1) =OLS-TT $=E(U_0|S=1)-E(U_0|S=0)$	-0.517[a] (0.183)	-0.222 (0.215)	-0.180 (0.483)	-0.376 (0.361)	-0.766[a] (0.315)	-0.639[b] (0.316)
Sorting Gain (S=1) =TT-ATE $=E(U_1-U_0|S=1)$	0.502[b] (0.240)	0.356[c] (0.258)	0.045 (0.266)	0.087 (0.261)	0.037 (0.158)	0.060 (0.142)
Selection Bias (S=0) = TUT – OLS $=E(U_1|S=0)-E(U_1|S=1)$	-0.104 (0.238)	-0.219 (0.200)	0.102 (0.446)	0.227 (0.324)	0.669[c] (0.433)	0.483[c] (0.314)
Sorting Gain (S=0) = ATE – TUT $=E(U_0-U_1|S=0)$	0.119[b] (0.059)	0.085[c] (0.061)	0.033 (0.196)	0.062 (0.188)	0.060 (0.250)	0.096 (0.226)
TT - TUT	0.622[b] (0.298)	0.441[c] (0.319)	0.077 (0.431)	0.149 (0.449)	0.097 (0.408)	0.156 (0.368)

Notes: Dependent variable is monthly wage in 1988, hourly wage in 1995 and 2002. OLS regressors are a binary 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 the propensity score as the instrument. The treatment effect estimates are based on results from local polynomial regression. Standard errors shown in parentheses are obtained by bootstrapping, and a superscript "a" denotes statistical significance level of 0.01, "b" of 0.05, and "c" of 0.10. All coefficients represent the estimated return to four years of college.

Figure 1.—Heterogeneous Returns and Instrumental Variable Method
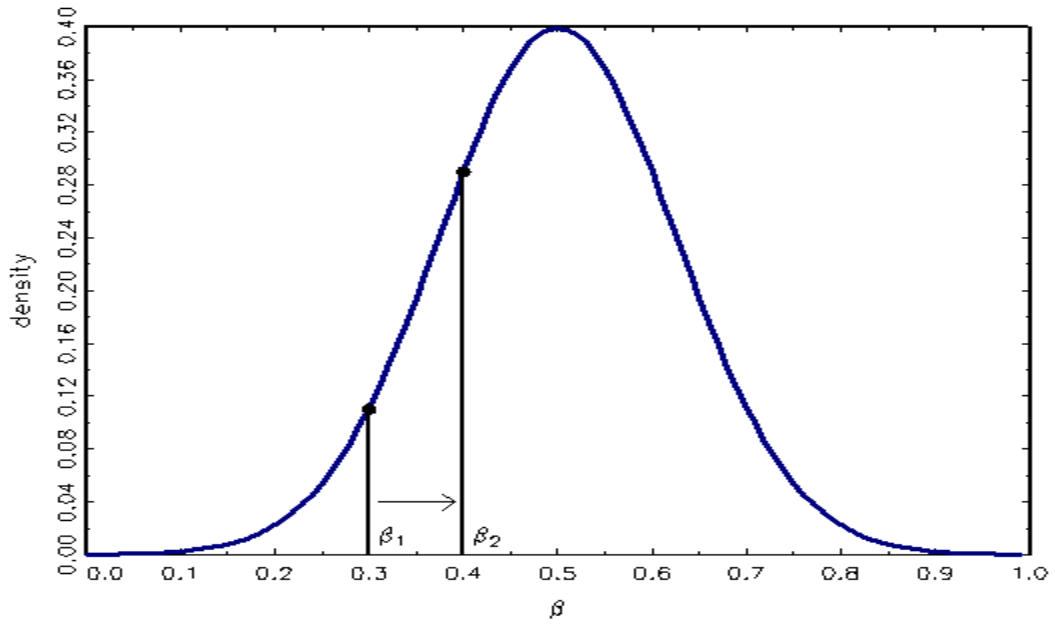
Figure 2.—Propensity Score Distribution: Kernel-Smoothed

41
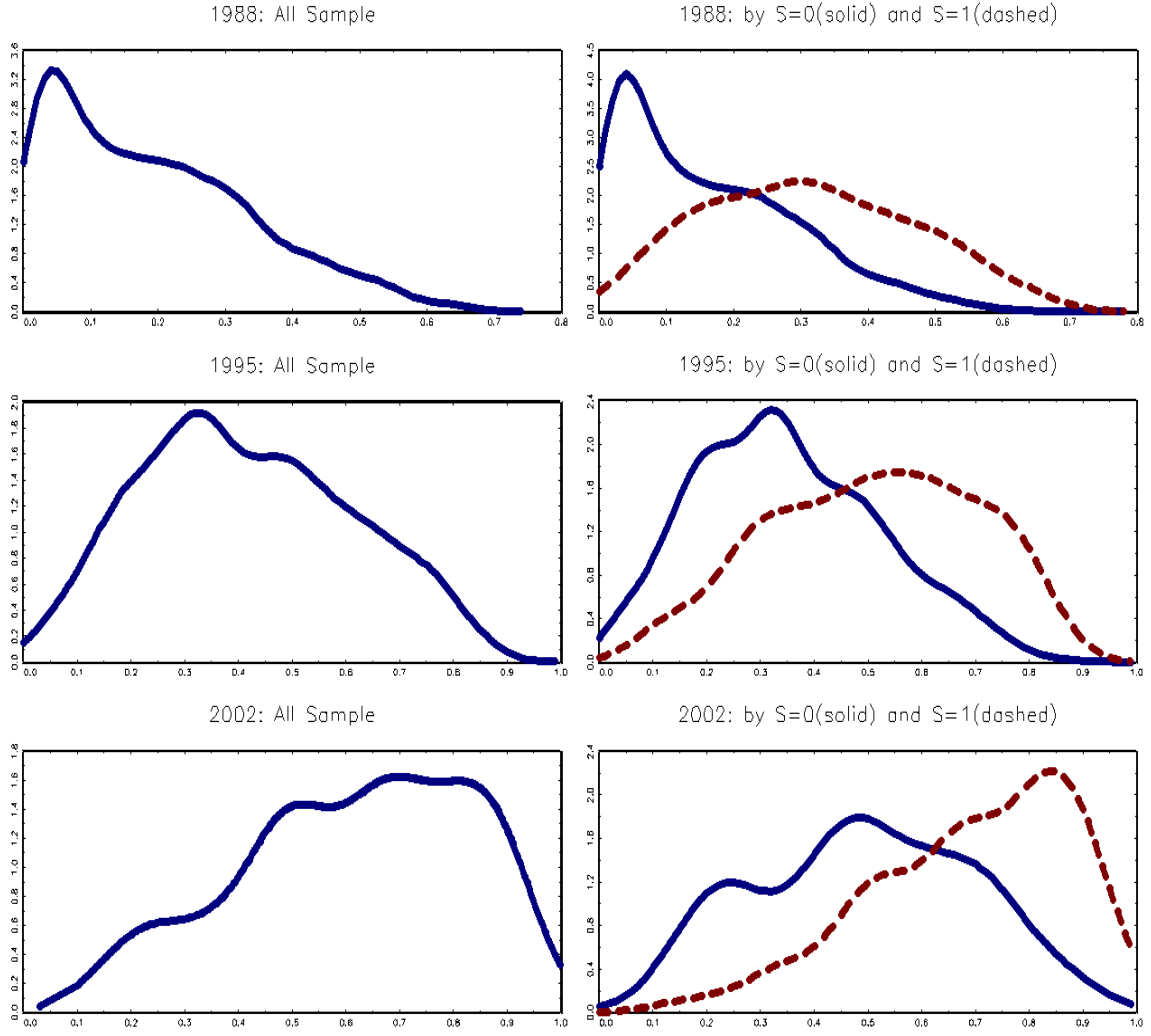
Figure 3.—Marginal Treatment Effects