Population Studies Center

Research Report

Yu Xie and Xiaogang Wu

Market Premium, Social Process, and Statistical Naivety: Further Evidence on Differential Returns to Education in Urban China

Report 05-578
July 2005

Population Studies Center
University of Michigan
Institute for Social Research

MARKET PREMIUM, SOCIAL PROCESS, AND STATISTICAL NAIVETY:

FURTHER EVIDENCE ON DIFFERENTIAL RETURNS TO EDUCATION IN URBAN CHINA*

Yu Xie

University of Michigan

Xiaogang Wu

Hong Kong University of Science & Technology

PSC Research Report 05-578

Direct all correspondence to Yu Xie, P. O. Box 1248, 426 Thompson Street, Population Studies Center, University of Michigan, Ann Arbor (Email: yuxie@umich.edu) or Xiaogang Wu, Social Science Division, Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong (Email: sowu@ust.hk). We thank Emily Greenman and Tony Perez for comments and suggestions.

Abstract

In an article published in the American Sociological Review (Wu and Xie, 2003), we addressed the substantive question of whether or not higher earnings returns to education in the market sector than in the state sector in reform-era China are caused by the market mechanism. Developing a typology of workers based on their labor market histories, we proposed a model of selective transition of workers from the state sector to the market sector as an alternative explanation for higher earnings inequality and higher earnings returns to education in the market sector. Our main empirical results showed that later market entrants have significantly higher returns to education than state sector stayers, whereas we found insufficient statistical power in the data with which to differentiate the education returns of early market entrants from those of either later entrants or stayers

In this paper, we address comments made by Jann (2005) about our article and his naive use of statistical tests in evaluating our findings. We reconceptualize the problem and illustrate the power of causal analysis using the propensity score method, which depends on the ignorability assumption. In considering two groups – early and late entrants to the market sector – our results suggest a strong selection mechanism at work. That is, when workers with a low latent propensity of making a transition indeed did make a late transition to the market sector, they benefited the most from the transition. In other words, the story is more about under what conditions workers migrated to the market sector than how workers benefit from migrating to the market sector.

Market Premium, Social Process, and Statistical Naivety: Further Evidence on Differential Returns to Education in Urban China

Statistics is a powerful, yet potentially dangerous, tool. More than two decades ago, the late Otis Dudley Duncan warned us of the danger of “statisticism”: “the notion that computing is synonymous with doing research, the naïve faith that statistics is a complete or sufficient basis for scientific methodology, the superstition that statistical formulas exist for evaluating such things as the relative merits of different substantive theories...” (Duncan 1984:226). For a long while, it was well understood that Duncan’s warning was directed toward quantitative researchers in sociology. Now it is clear that his concern is equally applicable to readers, even some careful readers, of quantitative research.

In his comment on our previous work (Wu and Xie 2003), Jann (2005) presents a poor understanding of the market transition process in China and the substantive question that we addressed — whether or not higher earnings returns to education in the market sector than in the state sector are caused by the market mechanism. As authors, we bear some responsibility for his misunderstanding of our work and will explain where he went astray. Further, we would like to take this opportunity to draw a general lesson for all of us: the limitation of relying on “canned” statistical tests and the benefit of basing statistical analyses on the substantive knowledge of social processes.

In the title of the original article (Wu and Xie 2003), we asked the question: “Does the market pay off?” Our emphasis was on the potential heterogeneity of workers in the market sector. Using work history data, we distinguished between two types of workers in the market sector: early birds and later entrants. We were concerned with the possibility that pooling early birds and later entrants, even if they each have the same education returns as stayers in the state sector, may make the returns to education appear higher in the market sector than those in the state sector.

Our study attempted to test two working hypotheses (Wu and Xie 2003:430): “Hypothesis 1: Earnings returns to education are higher for both later entrants and early birds than for stayers. Hypothesis 2: Later entrants, but not early birds, enjoy higher earnings returns to education than stayers.” Our main empirical results, which were confirmed by Jann, showed that later entrants have significantly higher returns to education than stayers, whereas there is insufficient statistical power in the data with which to differentiate the education returns of early birds from those of later entrants and those of stayers.[1] While the results do not support Hypothesis 1, they do not reject Hypothesis 2. We concluded that “commonly observed higher earnings returns to education in the market sector are limited only to recent market entrants” (Wu and Xie 2003:425).

[1] It was the first author who suggested to Jann that his problem with the Wu and Xie (2003) paper should be rephrased as a problem of insufficient statistical power.

Both the setup for our hypotheses and the results were merely descriptive, and were presented as such. We emphasized this point throughout the paper by alerting readers to the problem of aggregation when workers in a single group were in fact heterogeneous. While we questioned the prevailing wisdom that marketization *per se* "caused" the education returns to be higher, we never intended our statistical analysis to be more than a descriptive exercise. In such an exercise, formal statistical tests can be informative. However, when there is insufficient statistical power, statistical tests do not give researchers license to claim absurdity. In the original paper, we were modest in focusing on the contrast between later entrants and stayers – the only contrast that the empirical data afforded us to draw. But Jann wants more and wishes to overcome the data limitations by using better statistical tests. As a result, he would "draw the opposite conclusion and reject Hypothesis 2 in favor of Hypothesis 1," producing a contradiction that he then admits "make[s] no sense" (Jann 2005).

How does the use of seemingly correct statistical tests lead Jann so far astray? Why can we not blindly trust statistical formulas? This is a good lesson, especially for those who naively believe that, in Duncan's words, "computing is synonymous with doing research" and "statistics is a complete or sufficient basis for scientific methodology."

Jann errs in assuming that the three groups under discussion – stayers, early birds, and later entrants, are symmetric as would be observed in an experimental design. As a result, he borrows the language of multiple-group comparisons commonly used in ANOVA associated with experimental designs. However, we were dealing with observational data and were concerned with between-group and within-group heterogeneity generated by *social processes*. The earnings regimes for the three groups resulted from a *cumulative* social process in the past that is asymmetrical (Figure 1), and should be treated as such in an analysis. Comparing the three groups as if they were three experimental conditions is inappropriate and misleading.

Figure 1 presents a schematic flow chart of the respondent types in the 1996 survey of "Life Histories and Social Change in Contemporary China," the data used by Wu and Xie (2003). The Y-axis represents employment sector (state versus market), and the X-axis represents historical time. We make a convenient assumption that the market sector is an absorbing state so that there is no reverse transition from the market sector to the state sector.[2] In 1978, at the beginning of the Chinese economic reform, 1,197 respondents worked in the state sector. By 1987, 11 percent had made the transition into the market sector (d=1) and are called "early birds." Among the remaining 1,068 workers in the state sector and 522 new entrants who started in the sector between 1978 and 1987, 16 percent made the transition into the market sector (d=2) and are called "later entrants."[3] The remaining 1,337 respondents are called "stayers."

---

[2] That is, we excluded a small number of "market losers" from Figure 1 due to their small size (Wu and Xie 2003). We based the classification on the Wu and Xie's comprehensive measure of the market sector.

[3] 275 workers, including 82 "later entrants" in Wu and Xie's (2003) original article, who entered the labor force after 1987, are dropped from the analysis here.

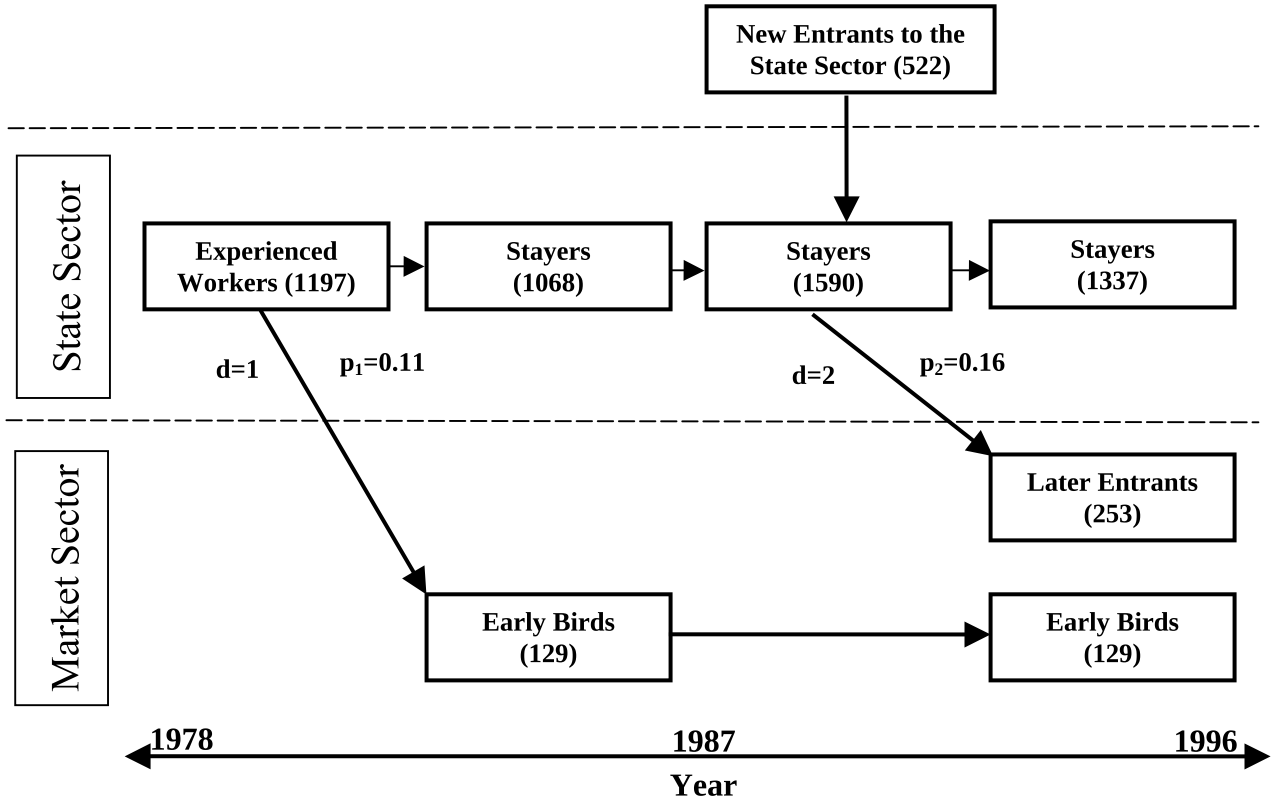

Figure 1. Flow Chart of Labor Market Transitions in China, 1978 – 1996.

We now reconceptualize the substantive problem in the language of causal inference with counterfactuals (Heckman 2005; Holland 1986; Manski 1995; Rubin 1974; Winship and Morgan 1999). Suppose we are interested in the causal impact of the entry to the market sector on (potential) future earnings in 1996. There are conceptually two questions in this setup: (1) what is the causal effect of an early transition (i.e., d=1), and (2) what is the causal effect of a late transition (i.e., d=2). There is an inherent asymmetry between the two. The second causal question is sensible only for those workers who did not experience an early transition. The first causal question involves the counterfactual comparison between those who experienced an early transition and those who did not, regardless what happened to them later. To borrow the notation for casual inference with time-varying treatments (Brand and Xie 2005), let Y_i^d denote the ith person's potential outcome if the person has made a transition at time d (d=1, 2, ∞), with d=∞ denoting that the person has not made a transition by the end of the study (i.e., a stayer). Note that for a person who has made an early transition (d=1), the counterfactual outcome should follow the principle of "forward looking sequential expectation" (Brand and Xie 2005) — a combination of a late transition (d=2) and staying (d=∞).

We thus define the average causal effect for the first question as:

$$
\begin{aligned}
&E(Y^{d=1}) - E(Y^{d>1}) \\
&= E(Y^{d=1}) - [E(Y^{d=2})P_2 + E(Y^{d=\infty})(1-P_2)].
\end{aligned}
\tag{1}
$$

Note that the transition probabilities are conditional so that $P_2 = P(d = 2 \mid d > 1)$. For the second question, the comparison is simpler, involving two regime-specific means:

$$E(Y^{d=2}) - E(Y^{d>2}) = E(Y^{d=2}) - E(Y^{d=\infty}) \tag{2}$$

It is never possible to compute quantities defined by equations (1) and (2), as we only observe one of the three potential outcomes for each worker. To infer the causal effects, it is necessary to introduce the ignorability assumption, which needs to be taken as provisional, as it is unlikely to hold in reality. The ignorability assumption states that all systematic differences associated with the transitions can be summarized by a set of observed covariates (X) (Rosenbaum and Rubin 1984).

Given this assumption, one may estimate the expected earnings based on the observed covariates, including education. This strategy was implicit in both Wu and Xie's (2003) and Jann's (2005) analyses. As shown in equations (1) and (2), we need four conditional expectations for the causal analyses: $E(Y^{d=1} \mid X)$, $E(Y^{d>1} \mid X)$ for the first question, and $E(Y^{d=2} \mid X)$, $E(Y^{d>2} \mid X)$ for the second question. The ignorability assumption means that $E(Y^{d=1} \mid X)$ can be estimated among early birds, $E(Y^{d=2} \mid X)$ among later entrants, and $E(Y^{d>2} \mid X)$ among stayers. However, $E(Y^{d>1} \mid X)$, being a weighted sum of two conditional expectations, should be estimated from both later entrants and stayers. Given that P_2 is small (at .16), a crude approximation of $E(Y^{d>1} \mid X)$ can be estimated from stayers (i.e., giving a full weight to stayers). This approximation is an interpretation of Wu and Xie's (2003) analysis strategy. Because later entrants constituted only a small proportion of the appropriate group against which to compare early birds, it makes little sense to compare, as Jann recommends, later entrants directly to early birds.

In the remainder of the paper, we illustrate the power of causal analysis using the propensity score method, which depends on the ignorability assumption (Rosenbaum and Rubin 1984). To borrow the jargon in the causal inference literature, we consider two "treatments" in our study, an early entry to the market sector and a late entry to the market sector. For the first treatment, the "control" group consists of workers who did not make an early entry and thus includes stayers as well as later entrants. For the second treatment, the "control" group consists of stayers only. The propensity score method allows us to summarize all the differences between the treatment and control groups in a single dimension — the probability of receiving a treatment. We then compute the average treatment effect on earnings within each propensity score stratum. A large literature shows that the propensity score method can remove large amounts of bias in causal inference with observational data (e.g., Dehejia and Wahba 1999; Winship and Morgan 1999). The propensity score approach is attractive in this case because it affords us an easy way to examine differences in observed covariates and compare the groups flexibly and non-parametrically.

In this framework for causal inference, the focus is explicitly on the cross-group differences in outcomes. Group differences in observed covariates (including education) are encompassed by the propensity scores. For the first transition, we estimate the propensity of P_1 via a logit model as a function of education, party membership, sex, seniority (including a square term), whether a family member was a cadre, rural registration status, parent's work status in the market sector when the respondent was age 14, province, interaction between party membership and education, and interaction between education and rural registration status. We find that human capital and political capital measures, such as education, party membership, seniority, and cadre connection, all predict negatively the probability of making an early transition to the market sector, confirming Wu and Xie's (2003:429) statement that "in early stages of economic reform entrants to the market sector tended to be those in the low tiers of the social hierarchy who were not at risk of losing privileges like those enjoyed by workers in the state sector."

We then group the respondents into 6 strata of estimated propensity scores to balance the distributions of both the estimated propensity score and the covariates between the treatment and comparison groups ($p<.05$). The number of cases in each stratum, separately by the treatment and control groups, is shown in Figure 2a.[4] The figure shows that early birds differ radically from other workers on observed covariates. Among workers who did not make an early transition, most have very low (0.28 or lower) propensities of making such a transition. In contrast, workers who did make the transition have relatively high estimated propensity scores. Without balancing these two groups through propensity score stratification, they would be incomparable.

We perform a similar analysis for the probability of making the second transition, i.e., P_2, conditional on a worker still being in the state sector by 1987. We include the following variables as predictors in the propensity score model: education (including a square term), seniority (including a square term and a cubic term), sex, party membership, and province. The propensity model for the second transition differs from that for the first transition because the mechanisms for making the transition changed.[5] For example, the influence of education now has an inverted-U shape, with a peak at about 3.5 years of schooling. In Figure 2b, we present the distributions of later entrants and stayers across 8 propensity score strata, again with the estimated conditional treatment probabilities and covariates balanced within each of stratum.

[4] The values on X-axis are the medians of the 6 propensity score strata. The 587 workers in the control group whose estimated propensity score is less than the minimum estimated propensity score for the treatment group are discarded.
[5] We fitted two propensity models specific to the two transitions to achieve balance within a propensity score stratum. The 753 workers in the control group whose estimated propensity score is less than the minimum estimated propensity score for the treatment group are discarded.

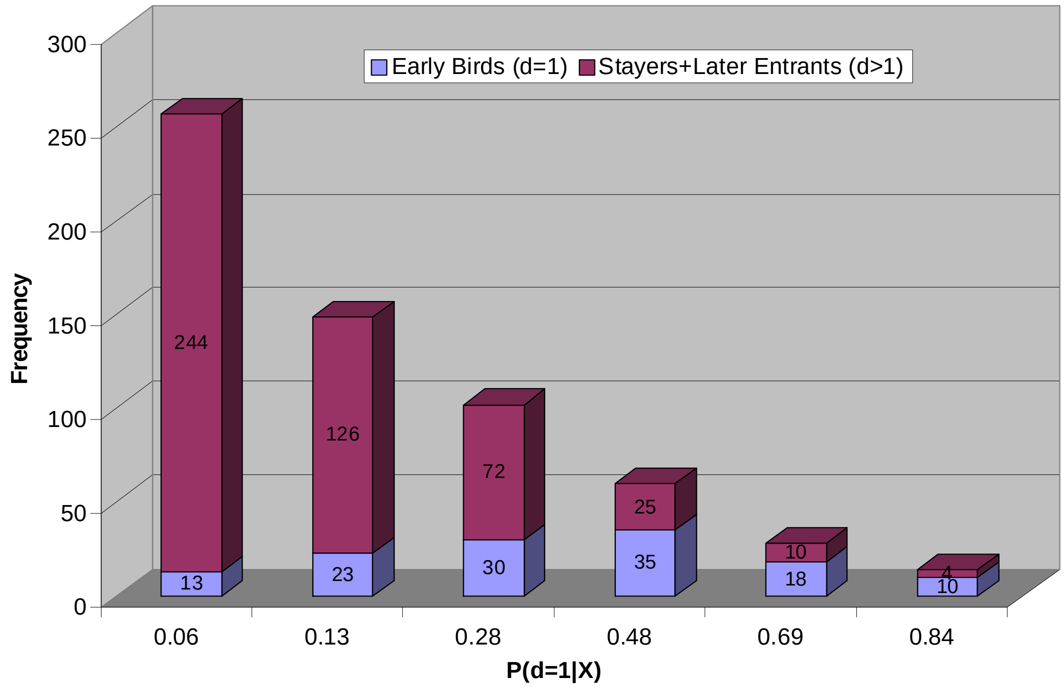

Figure 2a. Histogram of the Estimated Propensity Score for Early Birds (Treatment Group) and Stayers /Later Entrants (Control Group).

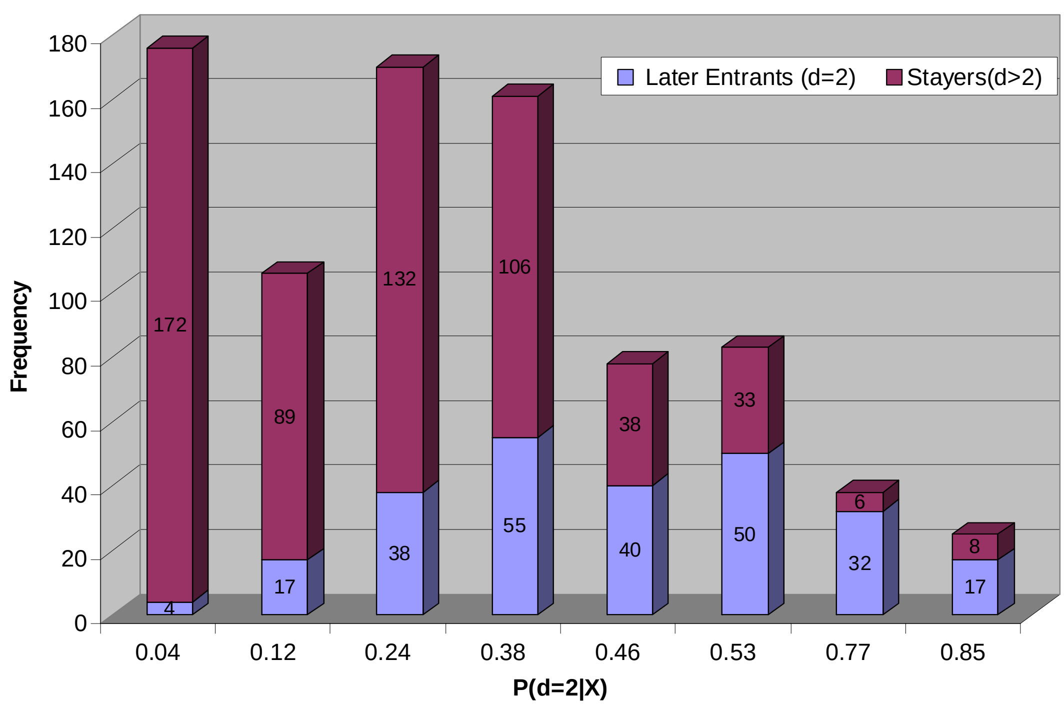

Figure 2b. Histogram of the Estimated Propensity Score for Later Entrants (Treatment Group) and Stayers (Control Group)

Under the ignorability assumption, there are no systematic differences between the treatment group (entrants to the market sector at a given time) and the control group (stayers in the state sector at a given time). The average earnings differences within a propensity score stratum can thus be interpreted as the average causal effect of market entry for that stratum. To answer the original question "does the market pay off?," we conduct detailed analyses of the treatment effect of market entry, in three separate steps. First, we estimate the treatment effects specific to propensity score strata. Second, we pool the results across strata under the assumption of a homogeneous treatment effect. Finally, we allow for heterogeneous treatment effects through a hierarchical linear model (HLM).

In Figure 3a we present summary findings from the analysis of treatment effects of an early transition. Dots represent point estimates of stratum-specific treatment effects, with corresponding t values (adjacent to the dots) for the null hypothesis that the treatment effect is zero. The results are clear. In no stratum can we detect a significant effect of an early transition. If we assume the treatment effect to be homogeneous across the strata and pool the estimates to improve precision, the overall treatment effect is estimated to be 166 RMB yuan, with a standard error of 97, resulting in an insignificant t value at 1.71. Finally, we allow the treatment effects to vary by strata in a hierarchical linear model (HLM) to examine whether the treatment effect varies systematically with the propensity score. The integer-score rank of a propensity score stratum is used as the predictor of the treatment effects across strata. The HLM results are represented by the linear line in Figure 3a. While it appears that the treatment effect increases positively with the propensity of being treated, this relationship is not statistically significant (t=1.51).

First, the treatment effect of making a late transition is relatively large and significantly different from zero for the four lowest propensity score strata. If we pool the different strata together for an overall treatment effect under the homogeneous effect assumption, the estimate is 236 RMB yuan with a standard error of 54, resulting in a highly significant t value of 4.36. However, the assumption of the homogeneous treatment effect seems to be violated, as there is clearly a downward trend in Figure 3b. The HLM model reveals that the size of the treatment effect strongly and negatively depends on the propensity score, with a unit change in stratum rank (i.e., crossing a propensity score stratum) associated with a reduction of 94 RMB yuan in the treatment effect (a significant relationship with t=-3.6). That is, the benefit of a late transition into the market sector is the greatest among those who were least likely to make the transition and diminishes with the propensity of making the transition.

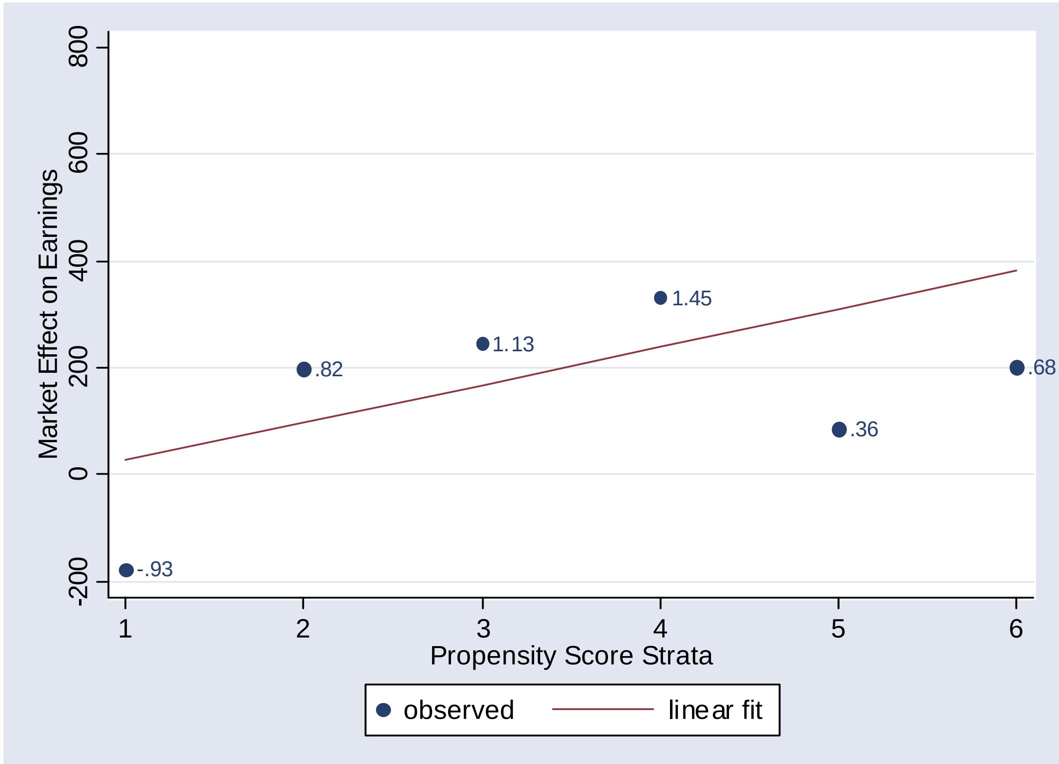

Figure. 3a. Market Treatment Effect on Earnings by Propensity Strata: Early Birds vs. Stayers and Later Entrants

Notes:

a. Numbers in the scatterplot are *t* values for earnings comparison between early birds (treatment group) and stayers/later entrants (control group). A *t* value less than 1.96 indicates there is no significant difference in earnings between the treatment and control groups within a propensity score stratum.

b. The linear plot is based on the Hierarchical Linear Model (HLM) estimates (level-2 model with slopes from level-1 model as outcomes regressed on propensity stratum rank). The effect of propensity stratum rank is not statistically significant (t=1.51).

Results for a late transition, summarized in Figure 3b, differ in several important ways. To the question "does the market pay off?" our results yield no simple answer. We do not find a generic market effect on earnings. The effect of the market on earnings varies, in two dimensions. First, confirming Wu and Xie's (2003) earlier results, we do not find a premium to an early transition to the market sector, while a late transition into the market sector is associated with higher earnings. Furthermore, we show that even among later entrants, the benefit of working in the market sector sharply decreases with the propensity of having made the transition. Hence, the summary finding of our reanalysis is that the market premium is only limited to late entrants who otherwise have a low likelihood of making a transition to the market sector. Who are they? A cost-benefit analysis would suggest that for a person to make such a transition, the benefit of market entry should exceed that of staying. Workers who do well in the state sector and are unlikely to

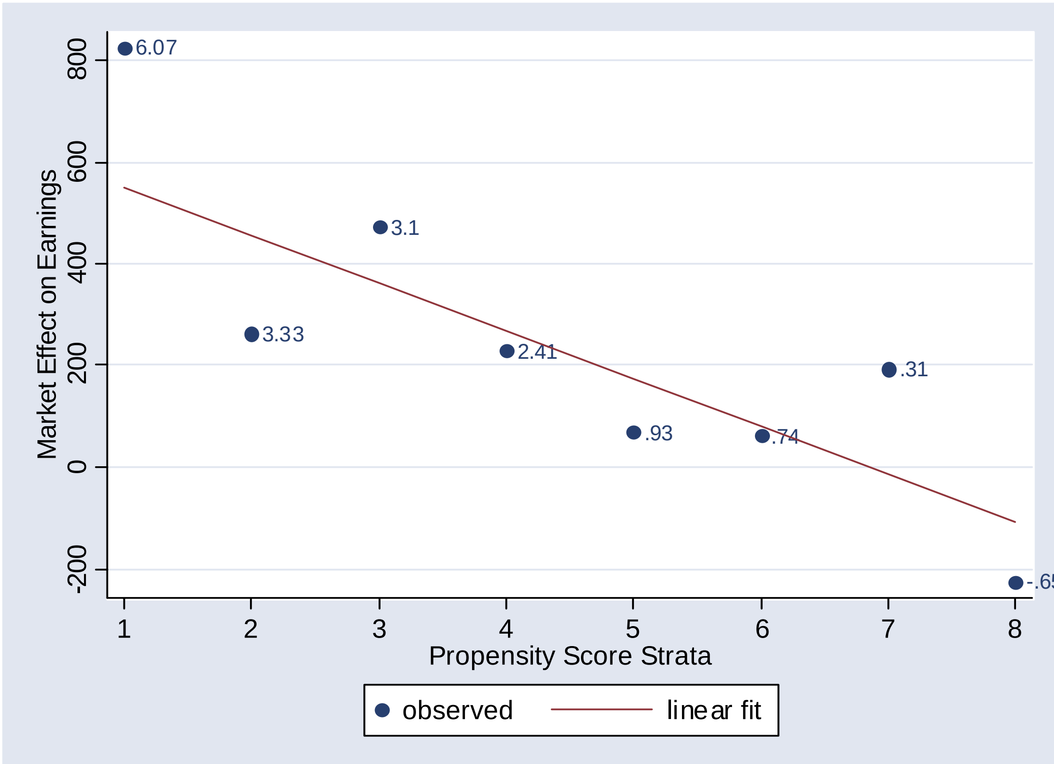

Figure 3b. Market Treatment Effect on Earnings by Propensity Strata: Later Entrants vs. Stayers

Notes:

a. Numbers in the scatterplot are *t* values for earnings comparison between late entrants (treatment group) and stayers (control group). A *t* value less than 1.96 indicates there is no significant difference in earnings between the treatment and control groups within a propensity score stratum.

b. The linear plot is based on the Hierarchical Linear Model (HLM) estimates (level-2 model with slopes from level-1 model as outcomes regressed on propensity stratum rank). The effect of propensity stratum rank is statistically significant (t=-3.6).

lose their jobs in the state sector have a low likelihood of entering the market. For them, the attraction of the market sector needs to be large enough to more than compensate for the advantages they already enjoy in the state sector.

In conclusion, our results suggest a strong selection mechanism at work: when workers with a low latent propensity of making a transition indeed did make a late transition to the market sector, they benefit the most from the transition. In other words, the story is more about under what conditions workers migrated to the market sector than how workers benefit from migrating to the market sector. During the economic reform in China, workers moved from the state sector to the market sector for different reasons, and not always voluntarily. That is the reason why they have also fared differently. If the market is such a monolithic, magic kingdom that always yields higher earnings than the state, we would not expect to see such wild variations in the earnings returns to entry into the market sector.

References

Brand, Jennie E. and Yu Xie. 2005. "Time-Varying Treatments, Time-Varying Effects: Causal Effects in a Longitudinal Setting." Annual Winter/Spring Meeting of Sociological Methodology Section, American Sociological Association in Chapel Hill, NC (April).

Dehejia, H. Rajeev and Sadek Wahba. 1999. "Causal Effects in Non-experimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of American Statistical Association* 94:1053-62.

Duncan, Otis Dudley. 1984. *Notes on Social Measurement, Historical and Critical.* New York, NY: Russell Sage Foundation.

Heckman, James J. 2005. "The Scientific Model of Causality." Unpublished manuscript. University of Chicago.

Holland, Paul W. 1986. "Statistics and Causal Inference" (with discussion) *Journal of American Statistical Association.* 81:945-70.

Jann, Ben. 2005. "Earnings Returns to Education in Urban China: A Note on Testing Difference Among Groups." *American Sociological Review.* This issue.

Manski, Charles. 1995. *Identification Problems in the Social Sciences.* Boston, MA: Harvard University Press.

Rosenbaum, Paul R. and Donald B. Rubin. 1984. "Reducing Bias in Observational Studies Using Sub-classification on the Propensity Score." *Journal of American Statistical Association* 79:516-24.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688-701.

Winship, Christopher and Stephen L. Morgan. 1999. "The Estimation of Causal Effects From Observational Data." *Annual Review of Sociology* 25:659-707.

Wu, Xiaogang and Yu Xie. 2003. "Does the Market Pay Off? Earnings Returns to Education in Urban China." *American Sociological Review* 68:425-42.

Population Studies Center Research Reports

Population Studies Center
University of Michigan
Institute for Social Research
PO Box 1248, Ann Arbor, MI 48106-1248 USA
http://www.psc.isr.umich.edu/pubs/