Earnings Returns to Education in Urban China: A Note on Testing Differences among Groups[*]

(Comment on Wu and Xie, *ASR*, June 2003)

Ben Jann

ETH Zurich

February 2005

Word count (excl. tables): 2400

Send correspondence to:
Ben Jann, Swiss Federal Institute of Technology Zurich, Department of Sociology, ETH Zentrum SEW E 28, CH-8092 Zürich, Switzerland; Phone: +41 1 632 55 58; Fax: +41 1 632 10 54; E-mail: jann@soz.gess.ethz.ch

[*] I would like to thank Elisabeth Coutts, Henriette Engelhardt, and Stefan Sacchi for their helpful comments and suggestions.

Biographical notes (word count: 93)

Ben Jann is employed in the Department of Sociology at the Swiss Federal Institute of Technology Zurich (ETHZ) and is a Ph.D. candidate at the University of Bern in Switzerland. His main areas of research are labor market inequality, environmental sociology, social research methods, and statistics. Recent publications include "The Use of New Technology and Occupational Mobility: An Event History Analysis of the Swiss Labor Market" (with Axel Franzen, *Journal of Applied Social Sciences Studies*, 2004), "Tabulation of Multiple Responses" (*The Stata Journal*, 2005), and *Einführung in die Statistik* (in German, Oldenbourg, 2005).

Earnings Returns to Education in Urban China: A Note on Testing Differences among Groups

(Comment on Wu and Xie, *ASR*, June 2003)

The finding that earnings returns to education are higher in the market sector than in the state sector in China's transition economy is commonly attributed to the "more efficiently operating market mechanisms" with respect to the valuation of human capital (Wu and Xie 2003, henceforth WX, p. 426). However, WX "challenge the prevailing wisdom that education is necessarily more highly rewarded in the market sector" (p. 425). Based on an empirical analysis of this issue, WX instead conclude that the market does *not* reward educational investment differentially.

Unfortunately, WX's conclusion appears premature. As we illustrate below, WX's conclusion is problematic mainly because of a general difficulty that exists in the interpretation of single pairwise tests in the analysis of the differences among three or more groups. The results of such comparisons can be ambiguous if the tests employed lack sufficient statistical power. WX compare the earnings returns to education among three different types of workers; however, they test only two out of the three possible group differences. Because the results of the omitted test are, in fact, at odds with the other results, WX's viewpoint is biased and leads them to an unjustified conclusion.

This comment is organized as follows. First, we briefly summarize WX's argumentation and hypotheses and then replicate their analysis in order to illustrate its shortcomings. We will then

derive an alternative answer to WX's research question while providing some general advice on how to address the problem of multi-group comparisons.

WX's Argumentation and Hypotheses

WX argue that the higher returns to education observed in the market sector could be due to selection effects in the "process of sorting workers into labor markets" (p. 426) and therefore propose a typology of workers based on individuals' labor-market histories. The worker types of interest are: (1) workers who have always been employed in the state sector ("stayers"), (2) workers who transferred to the nascent market sector early and stayed there ("early birds"), and (3) workers who were initially in the state sector but later switched to the market sector ("later entrants"). This typology is used to test competing hypotheses about the sources of observed differences between the two sectors in terms of returns to education.

If higher returns to education are caused by marketization per se, then both later entrants and early birds should enjoy higher returns than stayers, "because all workers in the market sector are subject to the same market mechanisms" (*Hypothesis 1,* p. 430). However, according to WX, potentially confounding factors were at work in China in the later stages of the economic transition. In particular, relatively skilled workers among the better educated were attracted to the market sector by large payoffs and relatively unskilled workers among the less educated were "pushed to the market sector through layoffs" (p. 430), an unobserved heterogeneity that would result in returns to education in the market sector appearing greater than they actually are. WX thus conclude that, if the market does not in fact pay off, the observable returns to education

should be higher in the market only for later entrants and not for early birds (*Hypothesis 2*). WX summarize this argument as follows:

> The crucial difference between the two hypotheses is the treatment of early birds. In Hypothesis 1, early birds are grouped with later entrants because they share the common feature of being in the market sector. (…) In Hypothesis 2, early birds are grouped with stayers because the two types of workers were approaching a convergence, against which later entrants were selectively recruited into the market sector. (P. 430)

WX also provide a graphical illustration of Hypothesis 2 (Figure 1:431). Note that the hypothesis depicted in their Figure 1, while not completely consistent with the description of Hypothesis 2 in the text body, makes a related prediction, namely that the apparently higher earnings returns to education in the market sector will disappear as soon the time of entry into the labor market is controlled for through the worker typology (early birds vs. later entrants). Either way, the basic idea is that the earnings returns to education in the market sector are biased (upwards) due to the heterogeneous composition of work force in the market sector and that accounting for the proposed worker typology in the analysis will, at least for some of the types, eliminate this bias.

Reanalysis and Criticism

In order to evaluate the Hypotheses 1 and 2 empirically, WX use data from a 1996 sample of the labor force in urban areas of China and estimate standard earnings regression models including

interaction terms for worker type and education.[1] The relevant part of the model can be represented as follows:

$$\log(\textit{Earnings}) = \beta_0 + \beta_1\, \textit{Education} + \ldots + \beta_6\, \textit{Later entrants} + \beta_7\, \textit{Early birds} + \beta_8\, \textit{Later entrants} \times \textit{Education} + \beta_9\, \textit{Early birds} \times \textit{Education} + \varepsilon$$

The stayers are the reference group, i.e. β_8 captures the difference between returns to education for later entrants and stayers and β_9 the difference between returns for early birds and stayers.[2] The Model 1 column in Table 1 contains WX's actual regression results (see Model 5c in their Table 4:437). Table 1 also contains a replication of their results, i.e. our own estimation based on the same data analyzed by WX (Model 2). The replication appears to be a reasonable approximation of the original estimates in Model 1.[3]

[Table 1 about here.]

[1] The data are from the 1996 survey on "Life Histories and Social Change in Contemporary China" and are obtainable from http://www.sscnet.ucla.edu/issr/da/lhsccs/chinaweb.html.

[2] The use of interaction terms might appear to complicate matters, but they are not relevant to the problem discussed in this comment. The essential issue for our purposes is that differences among more than two groups are examined, be it with conditional means or in conditional slopes.

[3] Note that the results in Table 1 are based on the comprehensive measure of the market sector, whereas WX also report results based on the restrictive and broad measures (for the definitions see WX:432). However, the use of a particular measure does not substantially alter the findings discussed here.

The main result of Model 2 (as well as Model 1) is that the effect of education differs significantly between later entrants and stayers, but not between early birds and stayers: β_8 is significant, but not β_9. WX thus conclude that one can distinguish between later entrants with higher observed returns to education on the one hand and stayers and early birds with a lower education effect on the other. Consequently, WX suggest that Hypothesis 1 should be rejected in favor of Hypothesis 2:

> Our findings suggest that the commonly observed higher earnings and higher returns to education in the market sector compared with the state sector in China are due entirely to the earnings outcomes of later entrants. Early market entrants resemble workers in the state sector in both their level of earnings and returns to education. Thus, it appears that it is not the market per se that renders higher rewards to later market entrants. Otherwise, early birds would enjoy an advantage similar to later entrants. (P. 438)

This interpretation does seem convincing at first sight. However, consider our Model 3 in Table 1, in which the reference group has been switched from stayers to later entrants. The primary result in this model is that the effect of education differs significantly between stayers and later entrants, but not between early birds and later entrants. Thus, if Model 3 is interpreted analogously to the manner in which Model 2 was interpreted, the conclusion is that there are higher earnings returns to education for later entrants and early birds, while the stayers have lower returns. Based on Model 3, one would therefore draw the opposite conclusion and reject Hypothesis 2 in favor of Hypothesis 1.

Models 2 and 3 are simply two different representations of exactly the same estimates. Such contradictory conclusions therefore make no sense. The reason for the seeming inconsistency is that both of the models test only a subset of group differences explicitly. For example, Model 2 tests later entrants and early birds against stayers, but does not test for differences between later entrants and early birds. One might argue that this is not a problem here because the two comparisons suffice to describe the relations among the three groups: If (1) A is different from C and (2) B is *not* different from C, then (3) A must be different from B (let A be the returns to education for later entrants, B those for early birds, and C those for stayers). Unfortunately, statistical tests do not allow this kind of inference. In the context of statistical testing, the true meaning of a statement like (2) is: "We do not know if B is different from C" (i.e., the hypothesis that B and C are equal cannot be rejected), and, thus, (3) does not follow from (1) and (2). Ergo, (3) must be tested explicitly.

We know from Model 3 that Assertion (3) cannot be confirmed for the data in question—an important piece of information that was neglected in WX's analysis; the assertion can also not be tested from the numbers provided in their paper. Thus, these data show neither that early birds have different returns to education than stayers nor that early birds have different returns than later entrants. In short, a lack of statistical power means that these results are inconclusive (a larger sample would eventually reveal a significant result for one or the other test and thus eliminate the inconsistency). There is therefore no justification for inferring that the observable earnings returns to education are higher *only* for later entrants. Actually, the population values for returns among later entrants and early birds could just as easily be the same, meaning that the

two groups could not be distinguished with respect to the effect of education. Together with the finding that the observed returns to education in the market sector as a whole are, in fact, significantly higher than those in the state sector, the effect of education could be higher for both market groups than the state group (despite the insignificant effect for early birds).

General Advice and Revised Conclusion

WX were led to a premature conclusion because they failed to systematically analyze all possible group differences. Such misinterpretations can be avoided by taking all possible contrasts into account (unless the pairwise differences are of no specific interest to begin with[4]). As illustrated above, one can do so by estimating a series of regression models in which the reference category is systematically varied. An equivalent but more efficient procedure is to perform so-called F-tests for linear restrictions based on the estimates of just one of the models (such tests make use of the variance-covariance matrix of the coefficients; the interested reader is directed to Greene 2003:95–99). The results of this procedure are displayed in Table 2, which contains values for the F-tests for all differences among the three groups in terms of earnings levels and/or returns to education. Again, it is apparent that any observed differences between early birds and later entrants do not approach statistical significance. Moreover, the p-values for comparisons

[4] Sometimes, polytomous variables are merely used as control variables. In such cases, the single coefficients associated with the variable may not be of particular interest—especially if the number of categories becomes large—and one would probably be satisfied with testing only the overall effect of the variable (see, e.g., Fox 1997:140–45). However, note that in any analysis a significant overall test is a prerequisite to a meaningful interpretation of the tests for single coefficients (in the present case, the overall test is significant).

between early birds and later entrants are of about the same size as the corresponding values for the comparisons between early birds and stayers. Taken together, these values indicate that a true difference between early birds and stayers is about as likely as a true difference between early birds and later entrants.

[Table 2 about here.]

Given these results, WX's Hypotheses 1 and 2 can be rejected with approximately equal likelihood. There therefore appears to be no empirical justification for favoring Hypothesis 2 over Hypothesis 1. If only one of these hypotheses is given credence, it should—at least if one adheres to the prescription of "Occam's razor" (Thorburn 1918)—be the more parsimonious of the two: Hypothesis 1. The revised conclusion is thus that the market *does* pay off, or at least that it is not possible to show that it does not with the data at hand. This finding is further supported by the fact that the higher earnings returns to education in the market sector as a whole remain practically unaffected by the separation of the main effects for early birds and later entrants (results not shown). Hence, the introduction of the worker typology proposed by WX does not seem to have any effect on the estimated earnings returns to education in the market sector.[5]

[5] Another significant problem with WX's analysis is that they use a dubious operationalization of worker types. The typology is based on employment in either the state sector or the market sector in 1987 and 1996. While respondents who were not active in the labor force in 1996 are excluded from the analysis, WX implicitly classify the respondents who were not active in 1987 (mostly respondents who had not yet begun their occupational careers) as belonging to the state sector in 1987. This procedure is problematic because the younger cohorts are coercively classified as stayers or later entrants. If the respondents who could not be classified for 1987 (i.e. respondents who were not professionally active in 1987) are dropped from analysis (253 out of 1539 stayers and 121 out of 332 later

References

Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods.* Thousand Oaks, CA: Sage.

Greene, William H. 2003. *Econometric Analysis.* 5th ed. Upper Saddle River, NJ: Pearson Education.

Thorburn, W. M. 1918. "The Myth of Occam's Razor" *Mind* 27:345–53.

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

---

entrants in our analysis), the difference in earnings returns to education between later entrants and early birds disappears almost entirely. This is further evidence against WX's Hypothesis 2.

Table 1. OLS Coefficients from Multiple Linear Regression of Monthly Earnings on Selected Variables, Urban China, 1996

	Model 1 *(Wu and Xie 2003)*	Model 2 *(replication of Wu and Xie 2003)*	Model 3 *(alternate specification of Model 2)*
Education (years of schooling)	.047***	.048***	.109***
	(.006)	(.007)	(.022)
Experience	.013***	.015***	.015***
	(.004)	(.004)	(.004)
$(\text{Experience})^2/1000$	-.167**	-.209**	-.209**
	(.060)	(.066)	(.066)
Party member (yes = 1)	.149***	.144***	.144***
	(.037)	(.037)	(.037)
Sex (male = 1)	.216***	.224***	.224***
	(.038)	(.038)	(.038)
Later entrants[a]	-.175	-.232	
	(.182)	(.183)	
Early birds[a]	-.067	-.035	.197
	(.249)	(.253)	(.256)
Stayers[a]			.232
			(.183)
Later entrants × Education	.056**	.061**	
	(.019)	(.020)	
Early birds × Education	.025	.026	-.035
	(.024)	(.027)	(.035)
Stayers × Education			-.061**
			(.020)
Constant	5.230***	5.185***	4.953***
	(.116)	(.111)	(.215)
Number of cases	2,060	2,057	2,057
R^2	.136	.142	.142

Source: Survey of "Life Histories and Social Change in Contemporary China" of adults aged 20 to 69, 1996.

Notes: Numbers in parentheses are robust standard errors adjusted for clustering. Data are weighted.

[a] Based on the comprehensive measure of the market sector; "stayers" are the reference category in Model 1 and 2, "later entrants" in Model 3; "market losers" (those who joined the market sector early and then left for the state sector) are omitted from the analysis because of the small number of cases they represent.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests)

Table 2. Hypothesis Tests for Differences among Early Birds, Stayers, and Later Entrants

Tests for Differences …	Null hypothesis	Observed value (standard error)	F-value (degrees of freedom)	p-value
… in earnings levels (based on a model without interaction terms):				
– later entrants vs. stayers	$\beta_6 = 0$	.289 (.072)	15.88 (1, 49)	.000
– early birds vs. stayers	$\beta_7 = 0$	.195 (.230)	.72 (1,49)	.400
– later entrants vs. early birds	$\beta_6 - \beta_7 = 0$	.094 (.181)	.27 (1, 49)	.606
…in returns to education (based on Model 2 in Table 1):				
– later entrants vs. stayers	$\beta_8 = 0$	.061 (.020)	8.88 (1, 49)	.004
– early birds vs. stayers	$\beta_9 = 0$	.026 (.027)	.96 (1, 49)	.333
– later entrants vs. early birds	$\beta_8 - \beta_9 = 0$	.035 (.035)	1.00 (1, 49)	.322
… in levels and returns (based on Model 2 in Table 1):				
– later entrants vs. stayers	$\beta_6 = 0$ $\beta_8 = 0$	-.232 (.183) .061 (.020)	11.23 (2, 48)	.000
– early birds vs. stayers	$\beta_7 = 0$ $\beta_9 = 0$	-.035 (.253) .026 (.027)	.57 (2, 48)	.570
– later entrants vs. early birds	$\beta_6 - \beta_7 = 0$ $\beta_8 - \beta_9 = 0$	-.197 (.256) .034 (.035)	.49 (2, 48)	.616

Source: Survey of "Life Histories and Social Change in Contemporary China" of adults aged 20 to 69, 1996.
Notes: The displayed tests are robust tests adjusted for clustering. Data are weighted.