Earnings differences between Chinese and Indian wage earners, 1987-2004: Trends and the impact of education*

Olivier Bargain
University College Dublin, and
IZA, Bonn
olivier.bargain@ucd.ie

Sumon Kumar Bhaumik
Brunel University,
William Davidson Institute, and
IZA, Bonn
Sumon.Bhaumik@brunel.ac.uk

Manisha Chakrabarty
Indian Institute of Management, Kolkata
m.chakrabarty@gmail.com

Zhong Zhao
IZA, Bonn
zhao@iza.org

August 9, 2007

Abstract

Chinese and Indian economies have expanded by over 50% in real terms over the 1990s. However, while there was a significant increase in the average earnings of wage earners in China during this period, the average earnings of their Indian counterparts remained stagnant. In this paper, using comparable earnings data for the two countries, and a common analytical framework, we address two empirical issues, the extent to which Indo-Chinese differences in average earnings can be explained by differences in endowment of and returns to education and other observed factors, and changes in earnings individually within China and India during 1987 to 2004. Our results suggest that while returns to education in China grew much more rapidly over this period than in India, this growth alone still does not explain satisfactorily the much faster rise in the average earnings of the Chinese wage earners. The increase in average wages in China is much more a combined product of an increase in average education levels and an increase in the returns to education.

Keywords: China, India, Earnings, Returns to education, Quantile Regression, Machado-Mata decomposition
JEL Categories: O15, J24, O53, P52

* The authors have benefited from conversations with Gary Fields and Jim Albrecht. They remain grateful to Klara Sabirianova and Yuriy Gorodnichenko for the Stata code on the Machado-Mata decomposition algorithm, and to the IZA and the CHIPS project for providing the Indian and Chinese data, respectively. Bargain and Bhaumik would like to acknowledge financial support and hospitality from IZA visitor program. The usual disclaimers apply. Corresponding author: Zhong Zhao, IZA, Schaumburg-Lippe-Str, 5–9, 53113 Bonn, Germany, +49-228-38 94 302, zhao@iza.org.

1. Introduction

Over the 1990s, responding to a range of market-oriented reforms, the Chinese and Indian economies expanded by over 50% in real terms.[1] Many aspects of this rapid growth – export growth, foreign direct investment, productivity growth, financial sector reforms, increasing spatial inequality, etc – have been explored in detail in the literature. However, much less has been said about the labour markets in these countries. While returns to education in both China and India have been estimated for various time periods,[2] with one notable exception (Kijima, 2006), no attempt has been made to examine what influenced earnings growth over time. Further, to our knowledge, there is no comparative analysis of issues related to Chinese and Indian labour markets. In this paper, we address this lacuna in the literature by examining in detail one specific empirical phenomenon, namely, that the average (real PPP-adjusted) earnings of Indian wage earners remained stagnant despite the rapid growth, while there was a significant increase in the average earnings of their Chinese counterparts.

To highlight this trend, in Figure 1 and Figure 2, we report the (weekly) earnings distributions for Indian and Chinese regular wage earners based on two national surveys of these two countries discussed in Section 2. In order to make the earnings comparable across time and countries, we convert them in 2000 USD PPP figures using a methodology described later in the paper. Figure 1 clearly reflects the faster earnings growth in China. In periods 1 (1987-88) and 2 (1993-95), the Indo-Chinese difference in earnings is positive for nearly all deciles of the earnings distribution. However, by period 3 (2002-04), the earnings gap had turned in favour of China for the lower half of the distribution, and was significantly reduced for the upper deciles.

Figure 2 confirms that earnings of wage earners in China were rising faster than those of their Indian counterparts. In period 1, a very significant proportion of the workers in the Chinese sample earned less than the median Indian worker. By period 2, however, while there are signs of a noticeable proportion of the Indian workers migrating towards the upper tail of the distribution, this migration is much more pronounced for the Chinese workers. Finally, by period 3, there are a smaller proportion of the Chinese wage earners in the lower tail of the distribution, compared with Indian wage earners, and,

---

[1] The GDP per capita for China was 1528, 2735 and 4568 in 1988, 1995 and 2002, respectively, measured in 2000 PPP US Dollar. The corresponding figures for India are 1569, 1994 and 2553.

[2] Papers in which returns to education were estimated include Byron and Manaloto (1990), Liu (1998), Knight and Song (2003), Zhang et al. (2005) and Appleton et al. (2005) for China, and Saha and Sarkar (1999), Kingdon and Unni (2001), Duraisamy (2002) and Kijima (2006) for India.

correspondingly, a greater percentage of Chinese workers in the middle earnings range than Indian workers. The Indians continue to dominate the upper tail of the distribution.

Figure 1: log-wage distribution (deciles)

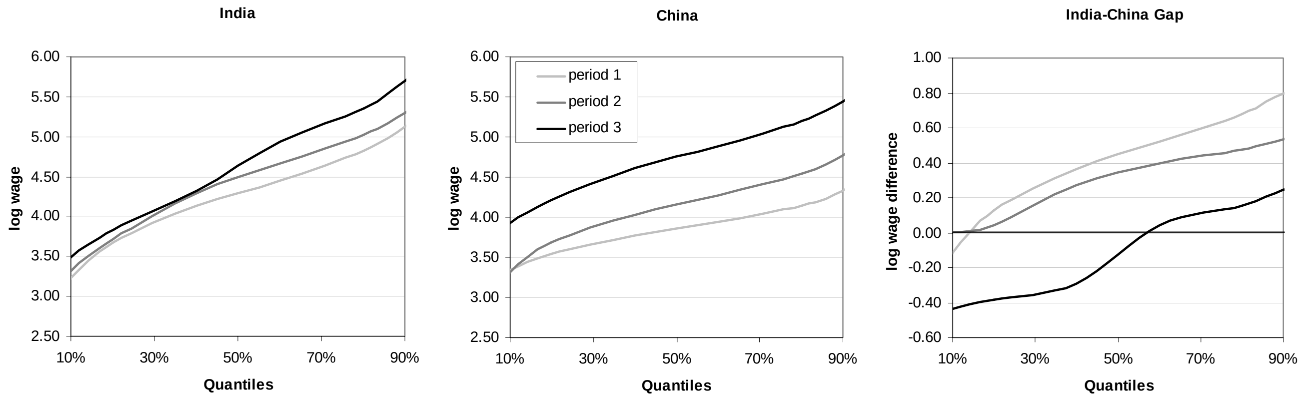
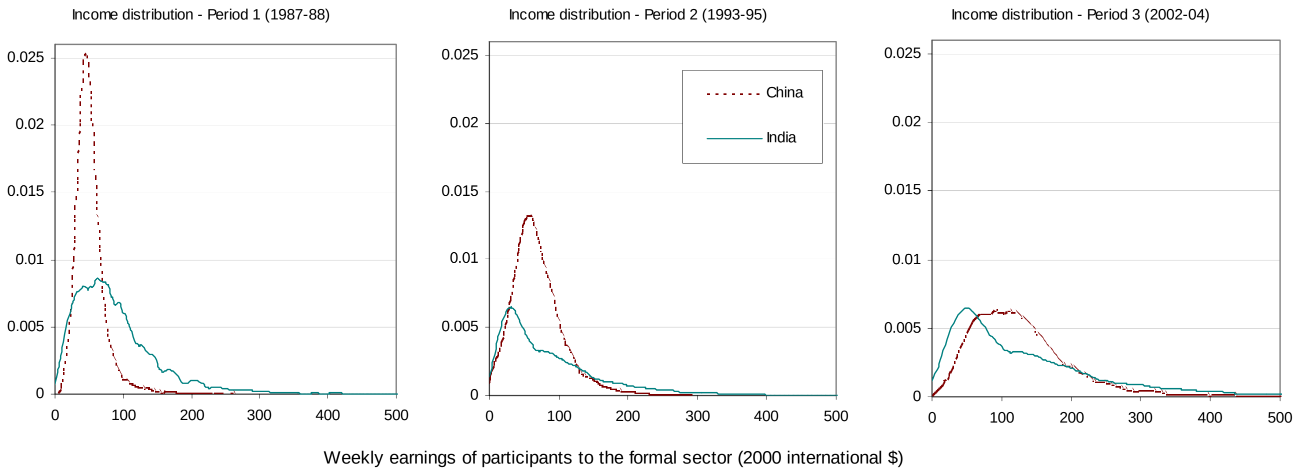


Figure 2: wage distribution (density)



The main objective of this paper is to shed light on the above mention earnings dynamics in China and India, and to explain the earnings differences between India and China. More specifically, in this paper, using comparable earnings data for the two countries, and a common analytical framework, we address two empirical issues. First, we examine the extent to which Indo-Chinese differences in average earnings can be explained by differences in endowment of and returns to education and other observed factors that affect earnings. Next, we take a look at the factors that might explain changes in earnings individually within China and India over this time-span.

In the course of our analysis, we use quantile regression models that are more suitable than ordinary least squares (OLS) for countries where heterogeneity within the labour force in terms of earnings and the impact of individual characteristics on earnings is significant.

Subsequently, we use the (decomposition) algorithm developed by Machado and Mata (2005) to examine the relative importance of differences in labourer characteristics and differences in the returns to these characteristics in explaining both Indo-Chinese earnings differences at a point in time, as well as differences in earnings across time within each country. The use of the Machado-Mata methodology allows us to study the coefficient effect at each quantile and account for heterogeneity in returns to individual characteristics (see Heckman and Li, 2004), as well heterogeneity in the characteristics themselves, across the earnings distribution. In particular, we focus on the role of differences in endowment of and returns to education that is perhaps the most important component of human capital in the stylised literature.

To reiterate, the Indo-Chinese gap in average earnings, which was in favour of the Indian wage earners in the late-1980s, declined rapidly over time, especially for the people in the upper earnings quantiles. Our results indicate that this shift cannot be totally explained in terms of a change in the relative returns to education alone. Returns to education rose significantly in China, especially between periods 1 and 2, and narrowed the gap with the corresponding measures for India, even though the returns in India continued to be higher in all three periods we considered. A rapid improvement in the average educational endowments of the Chinese wage earners, especially those at the lower end of the earnings distribution, played a more significant role in explaining the change in direction of the earnings gap in favour of China.

The rest of the paper is structured as follows. In section 2, we report the nature of the data and the associated descriptive statistics. Earnings estimations, and returns to education in particular, are presented and discussed in section 3. Section 4 reports the results of the decomposition analysis. Section 5 concludes and includes a brief discussion of policy implications.

2. Data description

Our empirical exercise is based on earnings data for regular wage earners in China and India. The data on the Indian wage earners are obtained from the 1987, 1993 and 2004 rounds of the National Sample Survey (NSS). These pan-Indian surveys are organised by the Central Statistical Organisation, and they use a stratified random sampling scheme to collect the data. The stratification is along geographical lines, with each state, as well as each district within a state, getting adequate representation (see Kijima, 2006). The Employment and Unemployment Schedule of the NSS is the only source of information for earnings and worker characteristics in India. It is stylised to exclude from the sample self-employed and casual workers, such that the sample includes wage earners who work full time and do not attend school. In addition, possibly to minimise measurement error, it is customary to restrict the sample to urban workers who account for more than 85% of the wage earners (see Kijima, 2006). We further restrict the sample to 21-60 year olds. After accounting for these adjustments, the size of the Indian sample was 22,480 in 1987, 21,681 in 1993, and 10,186 in 2004.

The Chinese data are obtained from the 1988, 1995 and 2002 waves of the China Household Income Project (CHIP).[3] Based on the large sample used by the National Bureau of Statistics, each of the three surveys gathers information from over 20,000 individuals, covering both rural and urban regions in eleven provinces in China and resembling the actual distribution of populations across these regions (Demurger et al., 2006). In order to make the Chinese sample comparable with the Indian sample, we restrict the former to urban wage earners as well, thereby making the Chinese sample similar to the one used by Liu (1998), Knight and Song (2003) and Zhang et al. (2005). After accounting for missing values and the aforementioned age restriction, the sample sizes for 1988, 1995 and 2002 are 16,519, 11,870 and 8,164, respectively.

Table 1: Descriptive statistics

	India			China		
Period	1	2	3	1	2	3
No of observations	22,480	21,681	10,186	16,519	11,870	8,164
Age	37.2	37.7	37.4	38.0	39.5	40.9
Female	0.15	0.16	0.18	0.47	0.48	0.44
No or primary education	0.33	0.26	0.20	0.12	0.06	0.02
Middle secondary education	0.13	0.14	0.16	0.38	0.30	0.21
High secondary education	0.29	0.31	0.34	0.36	0.40	0.40
College	0.24	0.29	0.30	0.14	0.24	0.37
Mean weekly earnings	89	104	140	52	70	134
Median weekly earnings	74	90	104	47	64	116

Period 1 is 1987 for India (1988 for China); period 2 is 1993/4 for India (1995 for China); period 3 is 2004 for India (2002 for China)
Selection: urban workers in formal sector, aged 21-60. Earnings are expressed in 2000 PPP international USD.

Both the surveys provide information on earnings, age, education and gender of labour force participants, the latter three individual characteristics being stylised determinants

[3] The CHIP project was jointly set up in 1987 by the Institute of Economics of the Chinese Academy of Social Sciences, the Asian Development Bank and the Ford Foundation; it also received support from the East Asian Institute of Columbia University.

of earnings. In order to make the earnings data comparable across the countries and the years, we make two important adjustments to the data. First, we concentrate on weekly earnings and transform all earnings into 2000 PPP US dollar equivalent, using the World Development Indicators on consumer price indices and PPP conversion factors. Next, given that the NSS/Indian survey includes information on the levels of education alone, while the CHIP/Chinese data includes both the number of years of education and the levels of education alone, we construct four education categories – no education or primary education, middle secondary education, high secondary education, and college education – that are comparable across the two countries. The use of education categories, while somewhat unusual in the Chinese context, can be found in other studies as well (e.g., Liu, 1998). Details about the construction of both the PPP-adjusted earnings measure and the comparable educational variables are available upon request.

The descriptive statistics for the CHIP/Chinese data and the NSS/Indian data are reported in Table 1. They indicate that in all three periods, our samples include Chinese and Indian wage earners of comparable age, with a Chinese worker being only marginally older than her Indian counterpart. Women constitute a significantly greater proportion of the wage earning work force in China than in India, perhaps reflecting higher educational attainment of an average Chinese woman,[4] as also the socialist ideology that promotes equal employment opportunity for men and women. Finally, it is evident that while India had an educational advantage in the late 1980s, the situation has changed rapidly since 1990s. In 1987-88, i.e., period 1 of our analysis, 24% of the Indian wage earners were college graduates, the proportion of people with middle secondary education or less being 46%. While the latter proportion for China was a comparable 50%, only 14% of the Chinese labour force had college education. By 2002-04, however, the picture had changed remarkably. The proportion of the Indian wage earners with middle secondary education or less declined to 36% and, correspondingly, there was a slight increase in the proportion of wage earners with college education, to 30%. Over the same period, the proportion of Chinese wage earners in the lowest two education categories more than halved to 23%, while the proportion of people with college education more than doubled to 37%. This rapid increase in the proportion of Chinese wage earners with college education is, at least in part, a manifestation of a bank-financed investment of RMB 200 billion (about USD 25 billion) in universities since 1998, which supplemented the government's budgetary support of about RMB 150 billion (USD 20

[4] Female literacy rate in China was 86% in 2002, compared with 48% in India in 2003.

billion) for secondary and higher education (2000 figures). For the sake of comparison, for the 1999-2000 fiscal year, the annual budget of the Indian government for secondary and higher education was INR 14.54 billion (USD 350-400 million, depending on exchange rate).

Table 1 also reports mean and median weekly earnings of Indian and Chinese wage earners during the three periods. It is easily seen that the compounded annual growth rate of average earnings was twice as large in China as in India between periods 1 and 2 (4.3% versus 2.2%), and three times larger between periods 2 and 3 (9.6% versus 3%). Median earnings in the two countries followed a similar pattern. In both countries, median earnings were lower than mean earnings, indicating the presence of large outliers in the upper tails of the earnings distributions.

Overall, the descriptive statistics indicate that (a) there was an increase in average educational attainment and average earnings of both Chinese and Indian wage earners, and (b) the rate of growth of both education and earnings was faster among the Chinese workers than among the Indian workers. In other words, it would be reasonable to hypothesise that the changes in the Indo-Chinese earnings gap over the years can be significantly explained by changes in the relative educational attainment of the Indian and Chinese wage earners. To the extent that a change in the educational distribution in a country is either a consequence of or results in (or both) a change in returns to education, changes in the relative returns to education might also explain the Indo-Chinese earnings gap to an extent. We revisit this issue later in the paper.

3. Earnings equation and rate of return to education

As mentioned earlier, our first endeavour is to estimate returns to education in China and India, using comparable data and specifications. Following the bulk of the literature, we estimate separate Mincer equations for workers in India and China. The model takes the following form:

$$\ln Y = \alpha_0 + \alpha_1 AGE + \alpha_2 AGE^2 + \sum_i \gamma_i EDUC_i + \sum_j \delta_j CONTROLS_j + \varepsilon \qquad (1)$$

where y is (weekly) earnings, age is a proxy for experience;[5] and *EDUC* is a vector of dummies capturing three different education levels (no or primary education is the omitted category). The control variables include gender (female = 1) and dummy variables for the

[5] Often, a proxy measure for experience is estimated by subtracting years of schooling and years prior to school enrolment (6-7 years) from age. However, since the Indian data do not provide information about years of schooling, we cannot undertake this exercise in this paper.

three youngest age cohorts (less than 23, 23-28 and 28-35).[6] Controlling for cohort effects helps to capture structural changes in the labour market over time and improve the fit of the model (Lemieux, 2006).

The Mincer equation is first estimated using ordinary least squares (OLS), which focuses on mean effects. In addition, we also perform quantile regression (Koenker and Bassett, 1978) to study the effects of covariates on earnings at different points of the conditional distribution. Specifically, we estimate quantile regressions for the 1st, 2nd, 4th, 6th, 8th and 9th deciles of the log-earnings distributions for each country. Note that we do not account for selection bias, i.e., the possibility that the workers in our sample did not become wage earners randomly but on account of some individual and household characteristics, in our estimation.[7]

Estimates of the Mincer equations for China and India are reported in Tables 2. We report both the OLS estimates as well as the quantile regression estimates for the two countries, for each of the three periods. It can be seen that the estimated coefficients are mostly significant, and the pseudo R-square values indicate a reasonable degree of fit of the Mincer specification given the cross-section nature of the data. Interestingly, the specification is much better at explaining inter-individual variation in earnings in India than in China.[8]

A look at the estimates for the coefficients of the control variables indicate that women in both countries earn lesser than their male counterparts, and gender discrimination, measured by the coefficient of the female dummy variable, is higher in lower earnings quantiles. Further, the extent of discrimination did not change significantly between the late

---

[6] A full breakdown of result by gender is available upon request to the authors. A specification where all variables vary with gender actually dominates the specification at use in the paper. However, given that our objective is to compare returns to education in China and India, as opposed to an analysis of gender differences in returns to education, we abstract from that analysis in this paper.

[7] To begin with, individual workers in developing countries with surplus labour often do not have the ability to rationally choose between forms of employment on the basis of relative marginal returns to their characteristics; choice of sectors and types of occupation is often accidental and driven by patterns of labour demand (Fields, 2005). Second, given that we have cross-section data for both countries for each time period, it is difficult to identify variables that can be used to identify a model with a selection equation and an earnings equation. Finally, despite an early attempt by Albrecht et al. (2006), extending the Machado-Mata decomposition methodology to account for selection is as yet not common practice; Machado-Mata remains the stylised methodology in the context of decomposing earnings gaps in heterogeneous samples (see, e.g., Nguyen et al., 2007)

[8] Zhang et al (2005) and Appleton et al. (2005) report higher goodness of fit measures for the Mincer specification for China. However, Zhang et al.'s specification includes controls for unobserved regional factors (i.e., dummy controls for regions), while Appleton et al. report significant impact of communist party membership on earnings, neither of which we can use in our specification for the sake of comparability between the specifications for China and India.

1980s and the earlier half of this decade. Not surprisingly, perhaps, gender discrimination is lower in China than in India.[9]

Table 2: Estimates of Mincer equation

Period 1

Coeff.	India OLS	India 10%	India 20%	India 40%	India 60%	India 80%	India 90%
age	0.10 ***	0.15 ***	0.13 ***	0.11 ***	0.09 ***	0.08 ***	0.05 ***
age2 /1000	-0.982 ***	-1.56 ***	-1.303 ***	-1.021 ***	-0.804 ***	-0.679 ***	-0.398 ***
female	-0.44 ***	-0.85 ***	-0.66 ***	-0.41 ***	-0.27 ***	-0.23 ***	-0.22 ***
mid-secondary	**0.22** ***	**0.24** ***	**0.26** ***	**0.22** ***	**0.20** ***	**0.15** ***	**0.14** ***
higher secondary	**0.57** ***	**0.63** ***	**0.63** ***	**0.56** ***	**0.52** ***	**0.47** ***	**0.47** ***
college	**0.97** ***	**1.04** ***	**1.04** ***	**0.96** ***	**0.92** ***	**0.90** ***	**0.89** ***
cohort1	0.07	0.11	0.11	0.09 **	0.06	0.06	-0.03
cohort2	0.10 ***	0.17 **	0.13 ***	0.13 ***	0.09 ***	0.07 **	-0.01
cohort3	0.05 **	0.04	0.06 **	0.08 ***	0.06 ***	0.05 ***	0.00
constant	1.58 ***	0.02	0.61 **	1.34 ***	2.00 ***	2.52 ***	3.24 ***
Pseudo-R2	0.37	0.21	0.22	0.23	0.24	0.25	0.25

Coeff.	China OLS	China 10%	China 20%	China 40%	China 60%	China 80%	China 90%
age	0.06 ***	0.07 ***	0.05 ***	0.04 ***	0.05 ***	0.06 ***	0.06 ***
age2 /1000	-0.547 ***	-0.624 ***	-0.43 ***	-0.364 ***	-0.388 ***	-0.567 ***	-0.572 ***
female	-0.13 ***	-0.16 ***	-0.12 ***	-0.10 ***	-0.10 ***	-0.10 ***	-0.14 ***
mid-secondary	**0.08** ***	**0.19** ***	**0.13** ***	**0.08** ***	**0.05** ***	**0.02** *	**-0.04** *
higher secondary	**0.13** ***	**0.26** ***	**0.18** ***	**0.13** ***	**0.08** ***	**0.07** ***	**0.01**
college	**0.22** ***	**0.35** ***	**0.25** ***	**0.20** ***	**0.16** ***	**0.15** ***	**0.16** ***
cohort1	-0.10 ***	-0.26 ***	-0.24 ***	-0.20 ***	-0.11 ***	0.03	0.04
cohort2	-0.08 ***	-0.16 ***	-0.15 ***	-0.11 ***	-0.09 ***	0.00	-0.01
cohort3	0.02 *	-0.01	-0.01	0.00	0.03 *	0.05 ***	0.04
constant	2.43 ***	1.80 ***	2.33 ***	2.63 ***	2.76 ***	2.64 ***	2.96 ***
Pseudo-R2	0.23	0.20	0.19	0.16	0.12	0.09	0.07

*Note: omitted variables are `no or primary education' and `cohort 4'. Stars report the level of significance (***:1%, **:5%, *:10%)*

Period 2

Coeff.	India OLS	India 10%	India 20%	India 40%	India 60%	India 80%	India 90%
age	0.10 ***	0.15 ***	0.15 ***	0.13 ***	0.10 ***	0.06 ***	0.06 ***
age2 /1000	-0.965 ***	-1.557 ***	-1.454 ***	-1.19 ***	-0.873 ***	-0.44 ***	-0.511 ***
female	-0.40 ***	-0.91 ***	-0.69 ***	-0.33 ***	-0.20 ***	-0.18 ***	-0.18 ***
mid-secondary	**0.27** ***	**0.30** ***	**0.26** ***	**0.27** ***	**0.21** ***	**0.21** ***	**0.17** ***
higher secondary	**0.58** ***	**0.56** ***	**0.62** ***	**0.60** ***	**0.53** ***	**0.52** ***	**0.49** ***
college	**0.96** ***	**1.00** ***	**1.07** ***	**1.03** ***	**0.95** ***	**0.94** ***	**0.93** ***
cohort1	0.10	0.21	0.22 **	0.18 ***	0.09 **	0.00	0.06
cohort2	0.10 **	0.14	0.18 ***	0.17 ***	0.08 **	0.02	0.05
cohort3	0.07 ***	0.08	0.12 ***	0.12 ***	0.07 ***	0.05 **	0.06 ***
constant	1.49 ***	-0.12	0.12	0.87 ***	1.80 ***	2.86 ***	2.96 ***
Pseudo-R2	0.26	0.15	0.19	0.23	0.25	0.25	0.26

Coeff.	China OLS	China 10%	China 20%	China 40%	China 60%	China 80%	China 90%
age	0.21 ***	0.39 ***	0.27 ***	0.15 ***	0.10 ***	0.09 ***	0.11 ***
age2 /1000	-2.368 ***	-4.56 ***	-3.061 ***	-1.552 ***	-0.982 ***	-0.942 ***	-1.145 ***
female	-0.23 ***	-0.28 ***	-0.19 ***	-0.15 ***	-0.13 ***	-0.12 ***	-0.14 ***
mid-secondary	**0.29** ***	**0.52** ***	**0.42** ***	**0.27** ***	**0.19** ***	**0.16** ***	**0.08** **
higher secondary	**0.50** ***	**0.78** ***	**0.66** ***	**0.44** ***	**0.33** ***	**0.27** ***	**0.22** ***
college	**0.68** ***	**1.00** ***	**0.84** ***	**0.61** ***	**0.44** ***	**0.41** ***	**0.34** ***
cohort1	0.55 ***	1.02 ***	0.69 ***	0.38 ***	0.20 ***	0.21 ***	0.34 ***
cohort2	0.28 ***	0.54 ***	0.40 ***	0.19 ***	0.05	0.15 ***	0.25 ***
cohort3	0.09 ***	0.20 ***	0.16 ***	0.07 **	0.00	0.04	0.05
constant	-0.63 **	-5.02 ***	-2.42 ***	0.44 *	1.81 ***	2.14 ***	2.08 ***
Pseudo-R2	0.16	0.16	0.13	0.10	0.08	0.06	0.05

*Note: omitted variables are `no or primary education' and `cohort 4'. Stars report the level of significance (***:1%, **:5%, *:10%)*

Period 3

Coeff.	India OLS	India 10%	India 20%	India 40%	India 60%	India 80%	India 90%
age	0.11 ***	0.14 ***	0.13 ***	0.11 ***	0.11 ***	0.08 ***	0.07 ***
age2 /1000	-0.962 ***	-1.373 ***	-1.209 ***	-0.92 ***	-0.908 ***	-0.565 ***	-0.552 ***
female	-0.41 ***	-0.75 ***	-0.65 ***	-0.49 ***	-0.28 ***	-0.16 ***	-0.16 ***
mid-secondary	**0.27** ***	**0.25** ***	**0.26** ***	**0.28** ***	**0.24** ***	**0.22** ***	**0.21** ***
higher secondary	**0.60** ***	**0.55** ***	**0.55** ***	**0.60** ***	**0.57** ***	**0.57** ***	**0.56** ***
college	**1.14** ***	**1.04** ***	**1.15** ***	**1.16** ***	**1.12** ***	**1.08** ***	**1.08** ***
cohort1	0.15 *	0.20	0.26 *	0.15	0.17 **	-0.02	-0.04
cohort2	0.08	0.11	0.09	0.07	0.10 *	-0.01	-0.02
cohort3	0.07 *	0.07	0.08	0.04	0.07 **	0.05	0.05
constant	1.40 ***	0.25	0.55	1.27 ***	1.54 ***	2.66 ***	3.02 ***
Pseudo-R2	0.42	0.20	0.22	0.26	0.28	0.27	0.26

Coeff.	China OLS	China 10%	China 20%	China 40%	China 60%	China 80%	China 90%
age	0.05 ***	0.09 ***	0.07 ***	0.03 **	0.04 ***	0.04 **	0.04 *
age2 /1000	-0.374 ***	-0.952 ***	-0.586 ***	-0.187	-0.202	-0.298	-0.279
female	-0.14 ***	-0.21 ***	-0.19 ***	-0.15 ***	-0.12 ***	-0.11 ***	-0.13 ***
mid-secondary	**0.21** ***	**0.19** **	**0.22** ***	**0.20** ***	**0.28** ***	**0.18** ***	**0.08**
higher secondary	**0.46** ***	**0.39** ***	**0.47** ***	**0.46** ***	**0.52** ***	**0.41** ***	**0.32** ***
college	**0.76** ***	**0.77** ***	**0.84** ***	**0.77** ***	**0.78** ***	**0.67** ***	**0.60** ***
cohort1	-0.03	-0.02	-0.03	-0.08	-0.01	0.08	0.01
cohort2	0.08	0.02	0.04	0.02	0.12 *	0.13	0.20 *
cohort3	0.03	0.01	0.04	0.00	0.07 *	0.04	0.00
constant	2.91 ***	1.42 **	2.13 ***	3.14 ***	3.19 ***	3.45 ***	3.88 ***
Pseudo-R2	0.17	0.10	0.11	0.11	0.10	0.08	0.08

*Note: omitted variables are `no or primary education' and `cohort 4'. Stars report the level of significance (***:1%, **:5%, *:10%)*

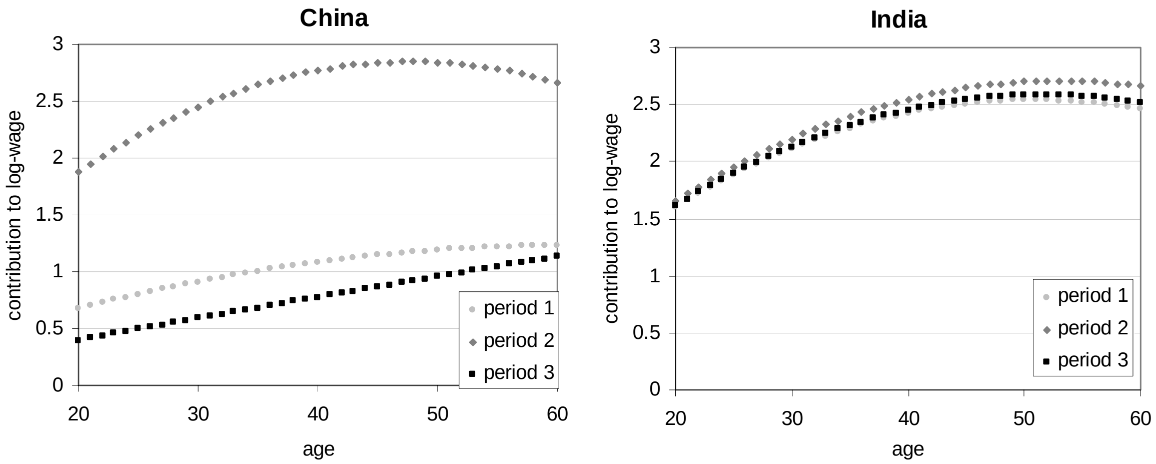
Estimates of the controls for cohort effects indicate that there were significant cohort effects in the first two periods, but not in the third. However, there were interesting inter-country and inter-period variations in this effect, Ceteris paribus, younger Indian cohorts

[9] Gender equality is known to greater in East Asia than in South Asia (King et al., 2000).

earned more than the older ones during both periods 1 and 2. In China, however, younger cohorts earned less, on average, than the older cohorts in the late 1980s, but this was reversed by the mid 1990s, and in period 2 the younger cohorts earned substantially more than their older counterparts. An explanation for the positive (younger) cohort effect can be found in Wu and Xie (2003). Their analysis suggests that in the 1990s, as market institutions got stronger in China, younger (and, often, new) workers were able to get private sector jobs that paid more than the public sector jobs in which people of older cohorts with similar education were employed. The negative cohort effect in period 1 is more difficult to explain.

The regression estimates also suggest that the age-earnings (or experience-earnings) profile in both countries is quadratic, with an inverted-U shape. As such, this is consistent with age/experience-earnings profiles observed in other countries. However, a closer examination of the coefficient estimates indicate that there are significant inter-country differences as well as differences in the age-earnings profile in China over time. In Figure 3, we report the age-earnings profiles generated from the OLS estimates reported in Table 2. The graphs indicate that return to age (or experience) was higher, on average, in India than in China, except in period 2, and that the return to age was much higher in China in period 2 than in the other periods.

Figure 3: wage progression



The reported change in the age-earnings profile in China between period 1 and period 2 was also observed by Appleton et al. (2005), who estimated the returns to experience to be the highest in 1995, followed by 1988, and with 2002 having the lowest returns to experience. They argue that the sharp decline in returns to experience after 1995 was possibly on account of an unnaturally high returns to this worker characteristic in the previous years such that when firms restructured after the mid 1990s the experienced workers were most likely to lose their jobs. This, in turn, put downward pressures on their employment, resulting in the decline in returns to education. They do not explain the rapid increase in these returns between 1988 and 1995. However, using the analysis of Knight and Song (2003), it is possible to infer that this may have been on account of the more experienced workers appropriating a greater than proportionate share of the (ostensibly performance-based) bonuses that were legitimised in the 1980s.

Our OLS estimates of returns to education are qualitatively consistent with earlier estimates for both countries (see footnote 2). The quantile regression results indicate the following: First, returns increase consistently with the education level, for both countries, for all time periods, and for all earnings quantiles. Second, with a few exceptions, the marginal impact of all education levels on earnings is lower for lower earnings quantiles than for higher earnings quantiles. This is especially true for China. Third, there were clear differences in the evolution of returns to education in China and India over time. In period 1, returns to education were higher in India than in China, particularly for higher secondary and college education. Returns to education rose rapidly in China between periods 1 and 2, catching up with, and in some cases, exceeding those in India in the latter period. Both the estimated low returns to education in China in the 1980s and the rapid increase during the early 1990s are consistent with the literature (Liu, 1998; Mauer-Fazio, 1999), as also with the stated policy of the Chinese government to closely link productivity and earnings of wage earners since the 1980s (see Knight and Song, 2003).[10] In period 3, returns to college education rose significantly in India, across earnings quantiles, while returns to other levels of education either remained the same or increased marginally. In China, on the other hand, returns to education rose, by and large, for all education levels for the upper earnings quantiles, but there was noticeable decline in returns to education among workers in the lower earnings quantiles. Significant rural-urban migration and the consequent undercutting of wages of those in the lower half of the earnings distribution is a plausible explanation for this decline in the returns to education of these wage earners.

4. Decomposition Results

Earlier in this paper, we noted that the Indo-Chinese earnings gap among wage earners, which was very much in India's favour in the late 1980s, narrowed significantly over time,

[10] Our estimates for returns to education are not directly comparable with those of Zhang et al. (2005) and Appleton et al. (2005); they used years of schooling as opposed to discrete educational categories.

and, by the early years of this decade, the gap had reversed in direction for the lower half of the earnings distributions. We also noted that there was a rapid improvement in the average educational endowment of Chinese wage earners during this time period, both in absolute terms and relative to their Indian counterparts. Finally, in the previous section, we noted that while returns to education in India remained higher than those in China, especially for higher education levels, returns in the latter country caught up significantly with those in India during the time period under consideration. Since endowment of and returns to education are among the most significant determinants of earnings, it is easy to argue that the narrowing (and, in some cases, reversal) of the Indo-Chinese earnings gap was significantly driven by the improved educational endowment and increasing returns to education of the Chinese wage earners. In this section, we examine the relative impact of changes in endowment and returns to education on the changes in earnings, using stylised decomposition methodology.

We undertake two distinct yet related decomposition exercises. First, we decompose the difference in average (log) earnings of Indian and Chinese wage earners into *endowment* (or *endowment*) *effect* and *coefficients effect*. The former measures the impact of the difference in the average characteristics (or endowments) of the Chinese and Indian workers on the differences in their earnings, while the latter measures the impact of the differences in the returns on these endowments on the same. Second, for each country, we similarly decompose the change in the average (log) earnings over time. In this case, the endowment effect would capture, for example, the impact of changes in the characteristics of Indian (Chinese) wage earners on the change in their average earnings across time. Similarly, the characteristics effect would capture the impact of changes in the returns to these characteristics of Indian (Chinese) wage earners on the same.

In keeping with the approach of Blinder (1973) and Oaxaca (1973), we can decompose the difference in (log) earnings, whether across countries or across time, in the following manner:

$$Ln\overline{Y}_I - \ln\overline{Y}_C \equiv \overline{X}_C'(\hat{\beta}_I - \hat{\beta}_C) + (\overline{X}_I - \overline{X}_C)'\hat{\beta}_I \quad (2)$$

where Y is (weekly) earnings, X is a vector of individual characteristics like age and education affecting earnings, β is a vector of returns to these characteristics, and subscripts I and C refer to India and China respectively. The decomposition of (log) earnings across time can be similarly characterized. The first term on the right hand side of the identity is the coefficients effect, i.e., the impact of differences in returns to the characteristics (between different periods, in this case) while average characteristics of the two samples are the same.

Correspondingly, the second term is the endowment effect, i.e., the impact of differences in mean characteristics of the two samples while the returns to these characteristics are the same across the samples.

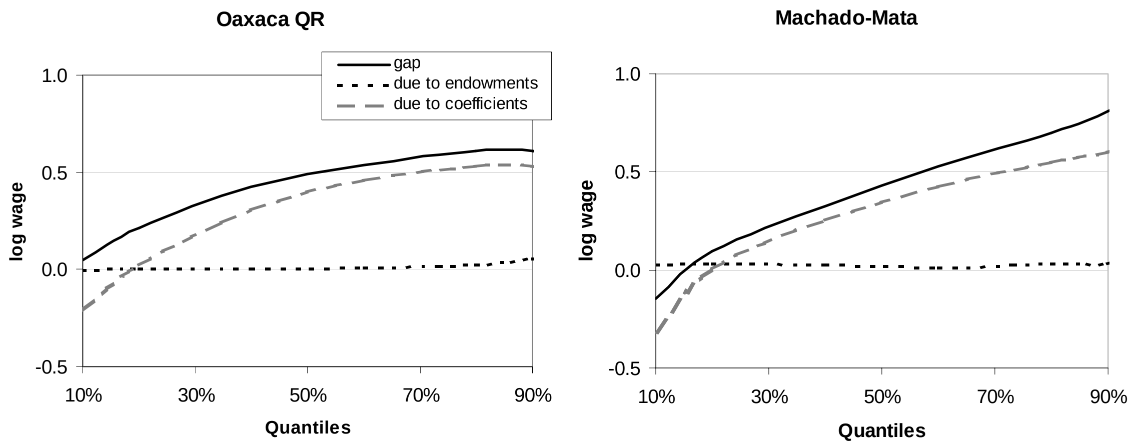
Classically, Blinder-Oaxaca decomposition is undertaken using sample mean of the characteristics and OLS coefficients that are estimates of the returns to these characteristics at the mean. However, since mean characteristics and (especially) returns to these characteristics can vary significantly across quantiles for a heterogeneous sample of individuals, it has become stylized in the literature to augment the classical decomposition methodology in two ways. To begin with, we can continue to use the mean values of the characteristics, but replace the OLS estimates with those generated by the quantile regression models. We call this the *Oaxaca-quantile regression* (or Oaxaca QR) method. Next, we use the Machado-Mata (2005) algorithm that decomposes the difference in (log) earnings in China and India at the n^{th} percentile of the distribution using coefficient estimates as well as values of the X vectors for the two countries at those percentiles. As demonstrated by Autor et al. (2005), the Machado-Mata approach nests most of the usual approaches. The method combines quantile regression and bootstrapping to generate two counterfactual density functions. Continuing with the example of decomposition of average (log) earnings across India and China, these counterfactual density functions are as follows: (i) the Chinese (log) earnings density function that would arise if Chinese wage earners had the same characteristics or endowments as their Indian counterparts, but continued to experience the same (or Chinese) returns to these endowments, and (ii) the density function that would arise if the Chinese wage earners retained their own characteristics but had the same returns to these characteristics as the Indian wage earners. A detailed description of the technique and an analysis of its asymptotic properties are provided by Albrecht et al. (2003) and Albrecht et al. (2006).

4.1. Decomposing wage differential between China and India

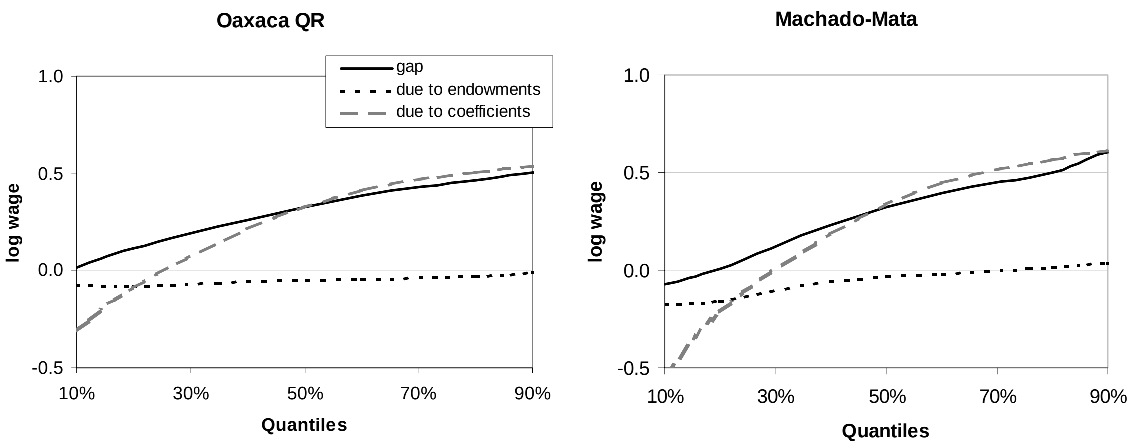
In Figure 4, we report the results of the decomposition for all three periods, obtained using both Oaxaca QR and Machado-Mata techniques. The results generated by the two algorithms highlight very similar patterns, indicating that our decomposition results are robust. The graphs reported in Figure 4 indicate that the coefficients effect explains most of the (log) earnings gap between the two countries, which is higher for higher earnings quantiles, at least for the two first periods. The endowments effect is close to zero, or even negative for all periods and for all earnings quantiles.

Figure 4: Oaxaca QR and Machado-Mata decompositions

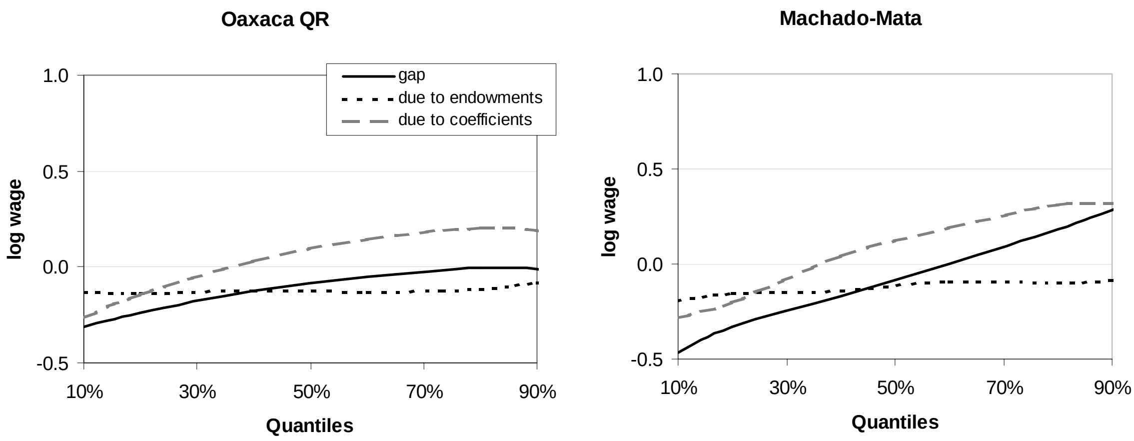
Period 1



Period 2



Period 3



Next, we examine the specific contribution of differences in endowment of and returns to education in China and India to the gap in (log) earnings between the two countries. In Table 3, we report the education-specific coefficient effects generated using the Oaxaca QR and Machado-Mata methodologies. It can be seen that these coefficient effects explain a significant proportion of the (log) earnings gap between the two countries for all quantiles. The effect is even more significant for the Machado-Mata estimates that are arguably more reliable than the Oaxaca QR estimates. Further, coefficient effect of education explains the earnings gap much more for the higher earnings quantiles than for the lower earnings quantiles.

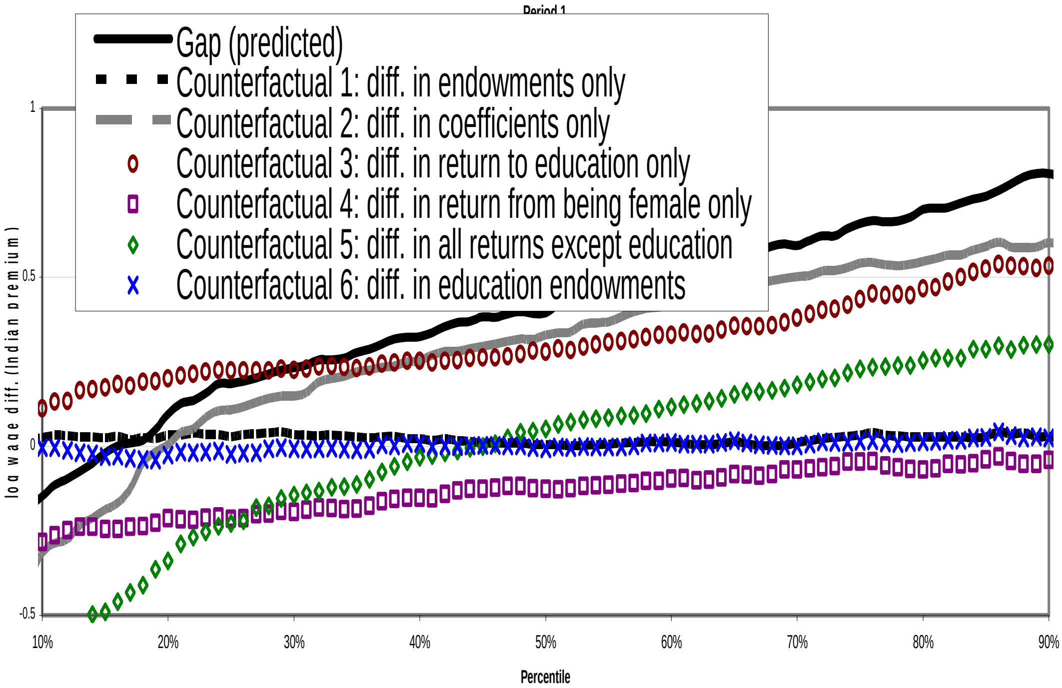
Table 3: India-China differential in returns to education

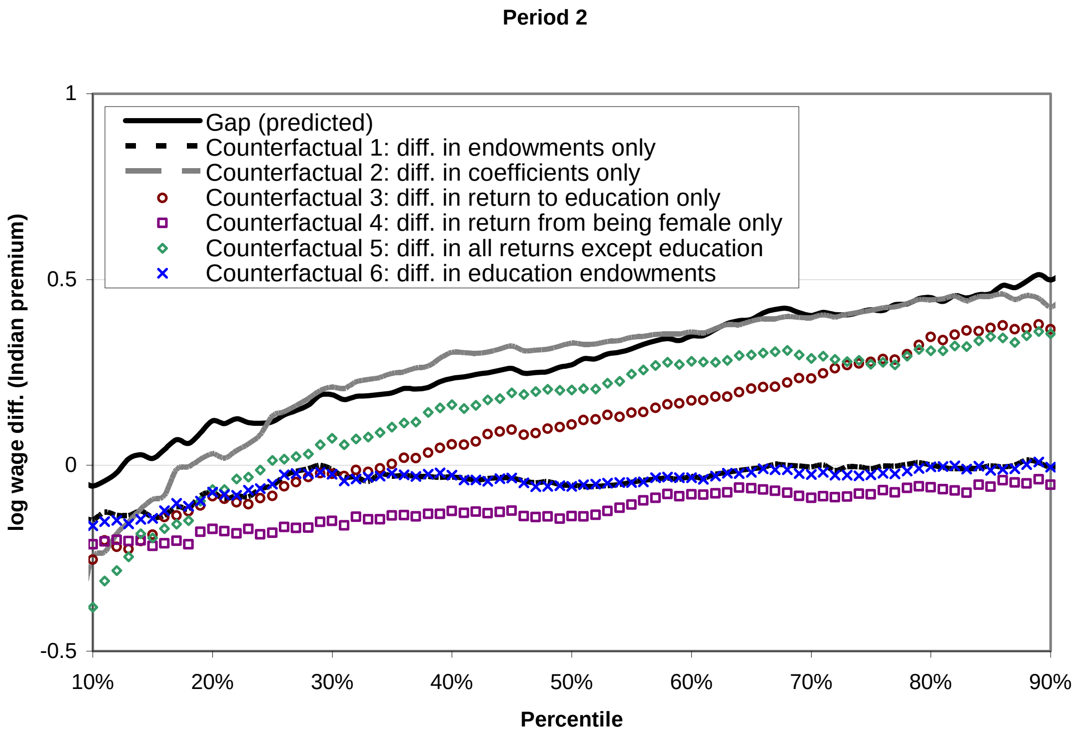
	period	10%	20%	40%	60%	80%	90%
Oaxaca QR coefficient effect							
	1	0.24	0.32	0.31	0.32	0.29	0.33
	2	-0.15	-0.01	0.16	0.21	0.24	0.28
	3	0.17	0.15	0.22	0.14	0.22	0.30
Machado-Mata counterfactual							
	1	0.11	0.20	0.25	0.33	0.47	0.53
	2	-0.25	-0.08	0.06	0.17	0.35	0.37
	3	0.08	0.12	0.12	0.16	0.26	0.31

For both techniques, figures represent the Indian–Chinese gap in log-wage which is explained by differences in returns to education

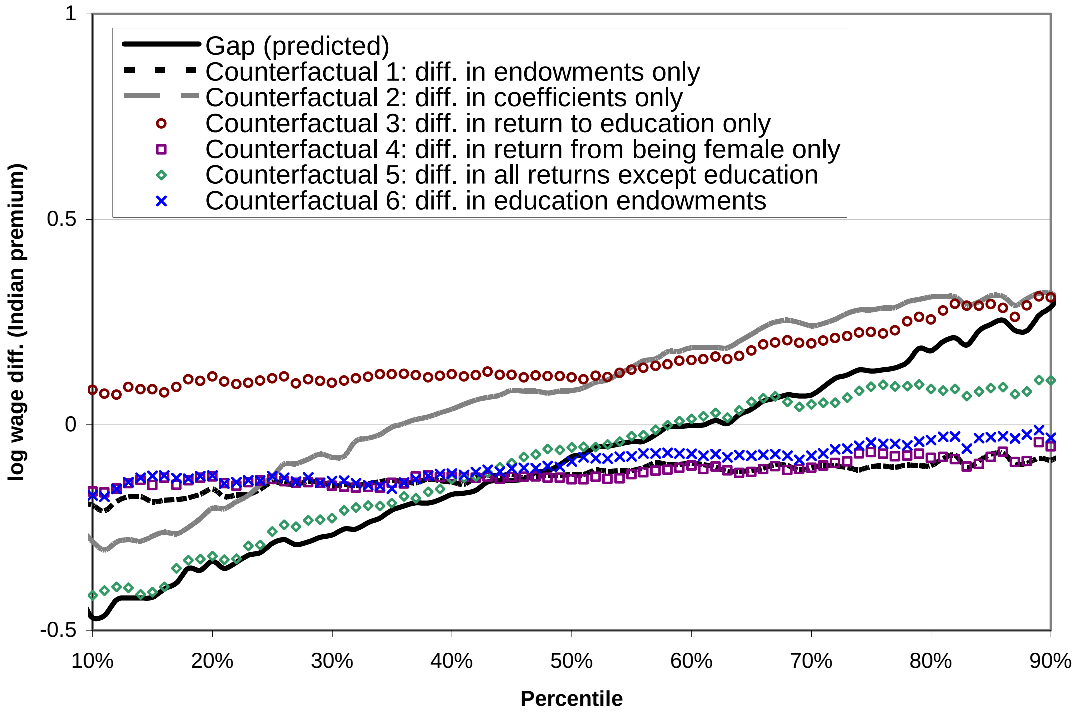
The characteristics (or endowment) effect and coefficient effect for education, generated using the Machado-Mata counterfactual, are reported in Figure 5. The graphs confirm that the differences in returns to education explain a significant proportion of the differences in (log) earnings at least during the first two periods. The role of differences in educational endowment in explaining the earnings gap is much less important, even though it accounts for most of the characteristics effect in our empirical exercise. Interestingly, in periods 1 and 2, the counterfactual for returns to education over-predicts the differences in (log) earnings for the lower earnings quantiles, and under-predicts it for the upper earnings quantiles. This is consistent with the earlier observation that the Indo-Chinese differences in returns to education are higher for upper earnings quantiles than for lower earnings quantiles. In period 3, this counterfactual over-predicts the earnings gap for all earnings quantiles, especially for the lower end of the distribution. In other words, earnings gap in period 3, which witnessed a decline in the advantage of the Indian wage earners, would have been higher, and in favour of the Indians, if the gap were determined by differences in returns to education alone. Once again, this is consistent with the higher estimated returns to education for India than for China.

Figure 5: Machado-Mata decomposition and differences in returns to education





Period 3

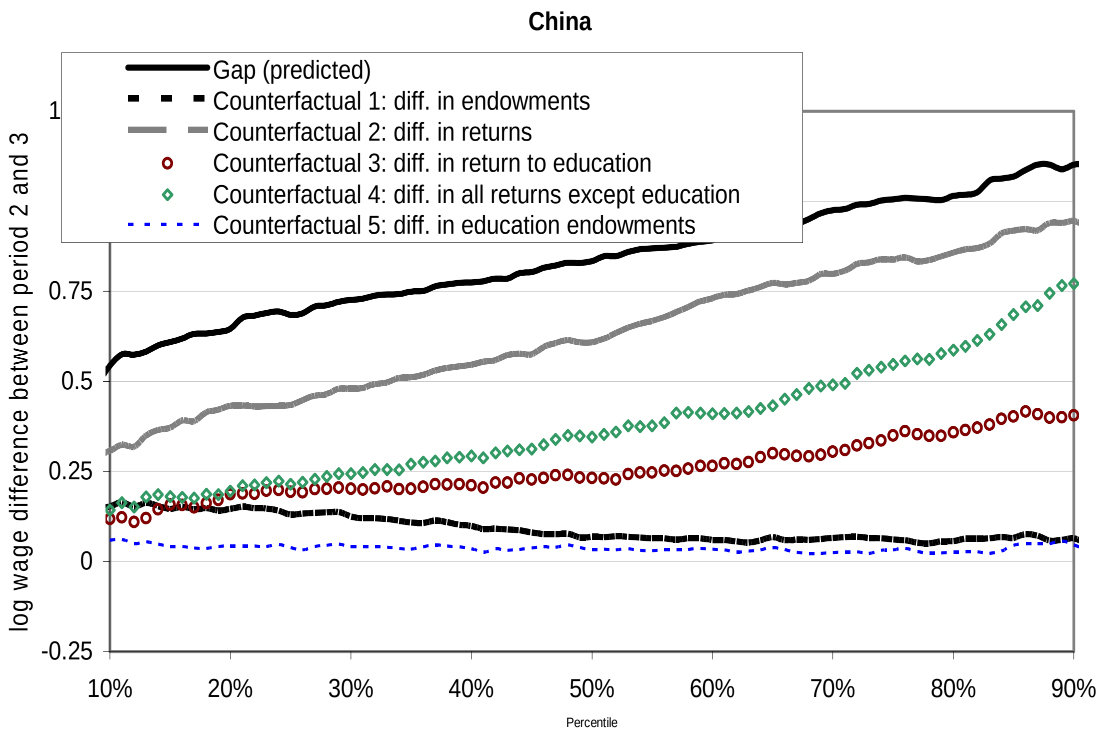


4.2. Decomposing wage changes across time within China and India

Next, we examine the determinants of the change in (log) earnings over time, within each country. Once again, using the Machado-Mata methodology, we decompose the differences in (log) earnings in each country between period 1 and period 3. The descriptive statistics and the coefficient estimates of the Mincer equation for the two countries, reported earlier in this paper, suggest the following: In the case of India, changes in returns to education and educational endowment of the wage earners would explain much of the change in average (log) earning between the two periods. However, as reported by Kijima (2006), the coefficient effect of education is likely to be much higher than the characteristic effect of this worker characteristic. In China, on the other hand, changes in both educational endowment of workers and returns to education are likely to play an important role in explaining the change in average (log) earning over time. While returns to experience may have played an important role in explaining earnings differences in China between periods 1 and 2, and periods 2 and 3, Figure 3 suggests that it is unlikely to explain much the earnings difference between period 1 and period 3.

The results of the Machado-Mata decomposition for (log) earnings difference between period 1 and period 3 are reported in Figure 6. In China, in keeping with expectations, the change (rise) in the return to education can explain about one-third of the change in (log) earnings between period 1 to period 3. Contrary to expectations, the endowment effect of education is small, and is more important for people at the lower tail of the earnings distribution. Overall, coefficient effect explains almost entirely the change in (log) earnings for all quantiles of the earnings distribution. Much of this coefficient effect can be traced to the change in the returns on unobserved characteristics that affect earnings, i.e., the difference in the constant terms. In India, where change in (log) earnings was much less for all earnings quantiles than in China, neither the characteristic effect nor the coefficient effect of education explains the change in average (log) earnings between period 1 and period 3 to a significant extent. Here too, coefficient effects of non-educational characteristics explain most of the change in average (log) earnings, but, unlike in China, the key driver of this non-educational coefficient effect is not differences in the constant term (see Table 2).

Figure 6: Machado-Mata decomposition between period 1 and period 3



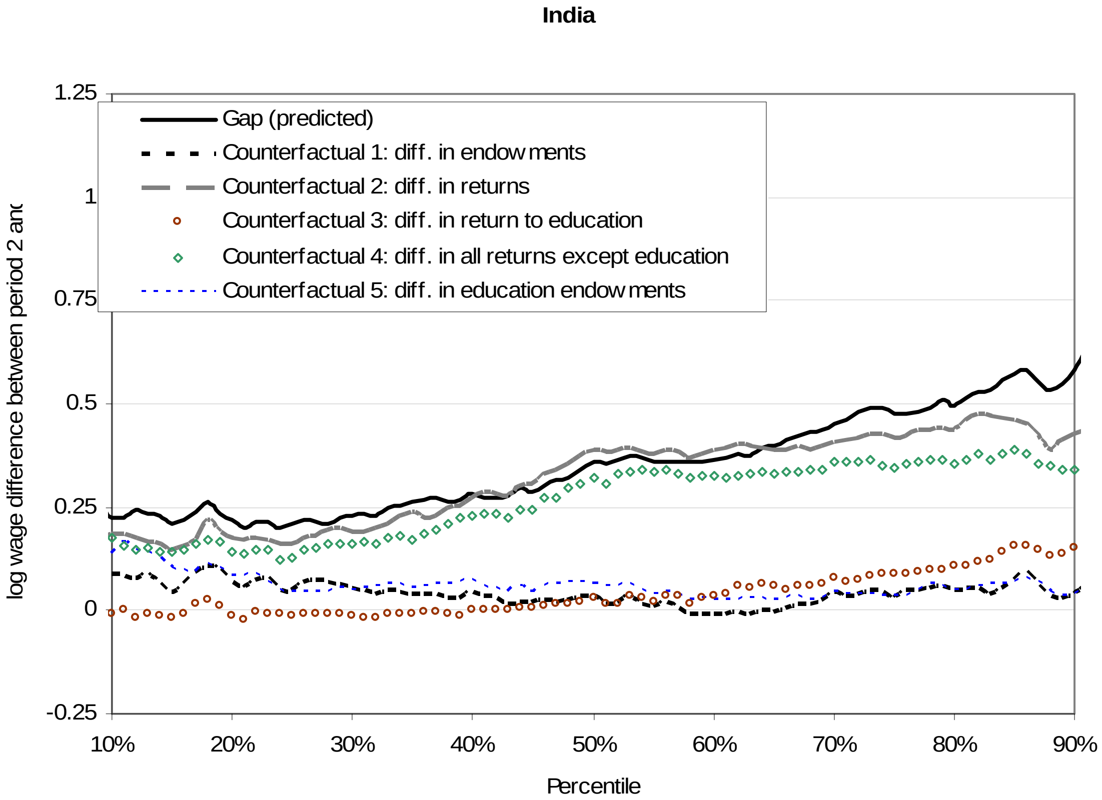
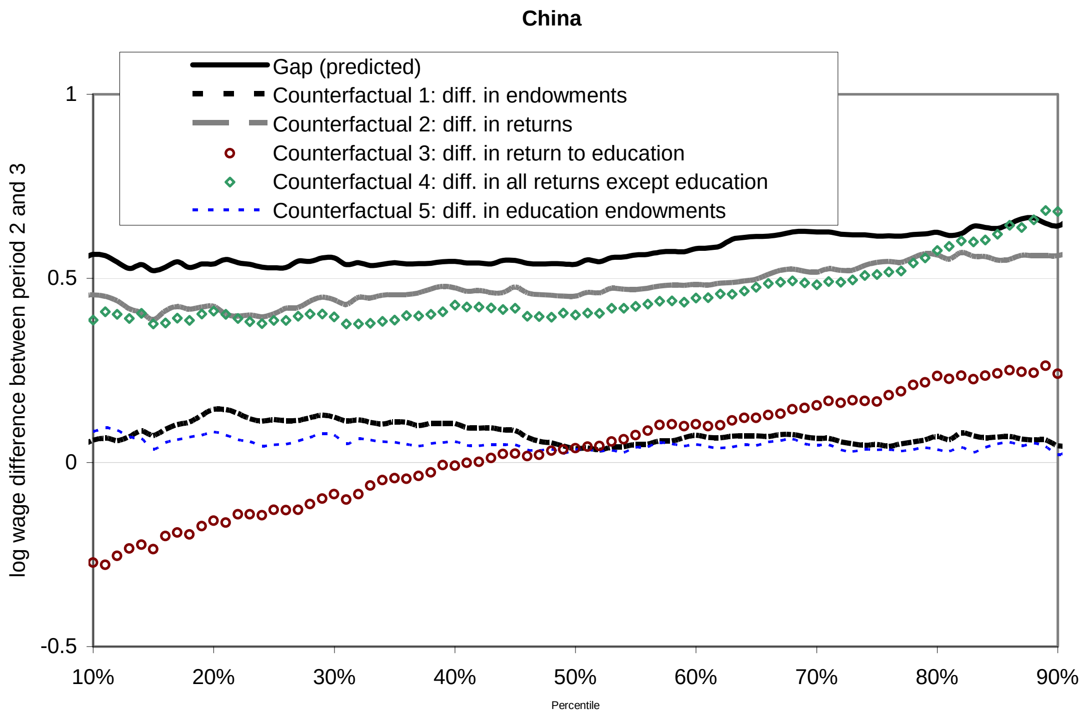
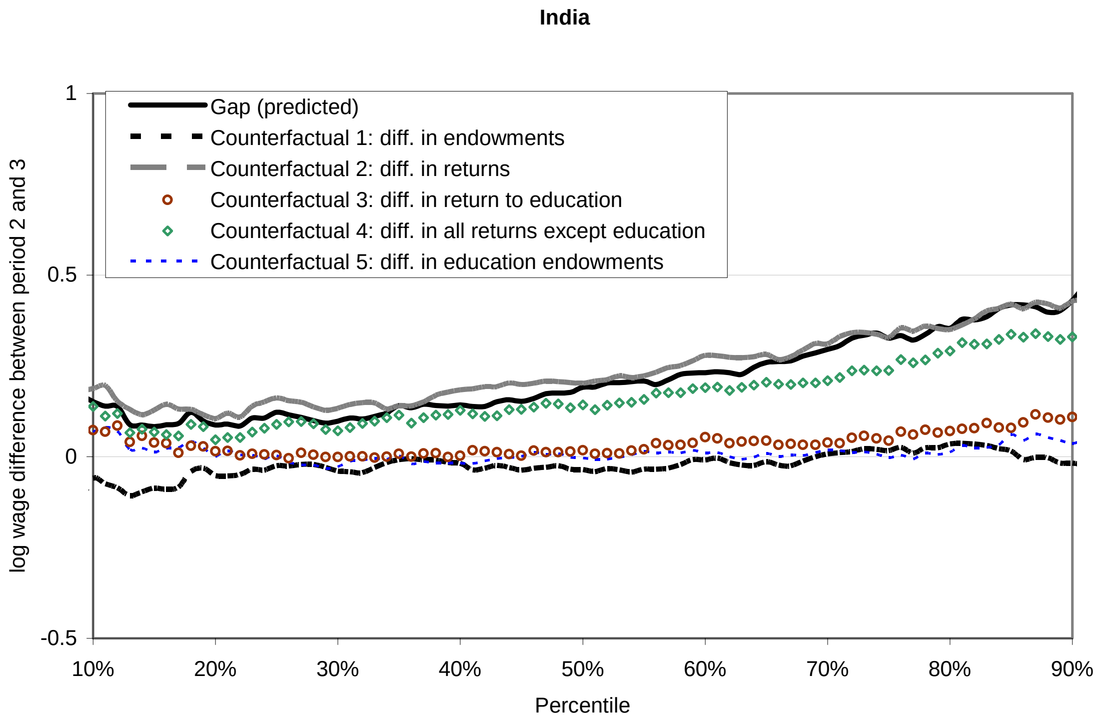


Figure 7: Machado-Mata decomposition between period 2 and period 3





There is one interesting observation for China: the earnings gaps between period 2 and period 3 are almost the same across income level, which is in sharp contract with period 1 to period 3. In another word, from period 2 to period 3, Chinese workers at different earning quantiles roughly benefit from the economic growth equally, while from period 1 to period 3, the higher income people benefit more.

The return to education can explain about one-forth to one-third of the earnings changes for high income people. But for lower income people, the earnings changes cannot be explained by return to education at all, in fact return to education should have a negative effect as shown in the graph, since we observe decreased return for this group of low income people. It seems for this group of people, the main force behind their wage increase is the improvement of their educational attainment.

5. Concluding remarks

Despite the near simultaneous rise of China and India as major economic powers, there are few rigorous comparative studies of the two countries. In this paper, we undertake a rigorous empirical analysis of one notable phenomenon that has been witnessed in both countries, namely, the steep rise in earnings of wage earners over time. The most interesting aspect of this phenomenon is that the rate of rise in earnings is much faster in China, a largely socialist (albeit rapidly reforming) country, than in India (which has had a mixed economy since independence, with a fairly large private sector). We particularly focus on the role of differences in educational endowment and returns to education in explaining both the inter-country differences in average (log) earnings during each of the three time periods being analysed, and also the change in average (log) earnings over time, within each country.

Descriptive statistics and quantile regression estimates indicate that the narrowing (and, in some cases, reversal) of the Indo-Chinese earnings gap, can in part be explained by rapidly rising educational endowment and returns to education in China, the rise in either being much more modest in India. Machado-Mata decomposition analysis indicate that differences in endowment and returns to education so indeed play a role in explaining the inter-country difference, but that the characteristic (or endowment) effect of education is much more modest than the coefficient effect. Coefficient effects in general play an important role in explaining differences in average (log) earnings for each country over time, but in the case of China the most important component of this coefficient effect is the difference in unobserved characteristics of the wage earners.

References

Albrecht, J., A. Bjorklund and S.C. Vroman, S. (2003) Is there a glass ceiling in Sweden? *Journal of Labor Economics*, **21**, 145-77.

Albrecht, J., A. van Vuuren and S.C. Vroman (2006) Counterfactual distribution with sample selection adjustments: econometric theory and an application to the Netherlands. Mimeo, Georgetown University, Washington, D.C.

Appleton, S., L. Song, Lina and Q. Xia, (2005) Has China crossed the river? The evolution of wage structure in urban China during reform and retrenchment. *Journal of Comparative Economics*, **33**:4, 644-663.

Autor, D, L. Katz and M. Kearney (2005) Rising wage inequality: the role of composition and prices, Working paper no. 11628, National Bureau of Economic Research, Cambridge, Massachusetts.

Blinder, A. S. (1973) Wage discrimination: Reduced form and structural variables. *Journal of Human Resources*, **8**, 436-455.

Byron, R.P. and E.Q. Manaloto (1990) Returns to education in China. *Economic Development and Cultural Change*, **38**(4), 783-796.

Demurger, S., M. Fournier, L. Shi and W. Zhong (2006) Economic liberalisation with rising segmentation on China's urban labour market. Working paper no. WP 2006-46, ECINEQ – Society for the Study of Economic Inequality, Spain.

Duraisamy, P. (2002) Changes in returns to education in India, 1983-1994: By gender, age cohort and location. *Economics of Education Review*, **21**(6), 609-622.

Fields, G. (2005) A Guide to Multisector Labor Market Models, World Bank working paper.

Gajwani, K., R. Kanbur and X. Zhang (2006) Patterns of Spatial Convergence and Divergence in India and China, Paper prepared for the Annual Bank Conference on Development Economics (ABCDE), St. Petersburg, January 18-19, 2006.

Heckman, J. J. and X. Li (2004) Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000. *Pacific Economic Review*, **9**:3, 155-171.

Kijima, Y. (2006) Why did wage inequality increase? Evidence from urban India 1983-99. *Journal of Development Economics*, **81**:1, 97-117.

Kingdon, G.G. and J. Unni (2001) Education and women's labour market outcomes in India. *Education Economics*, **9**(2), 173-195.

King, E.M. et al. (2000) Engender development though gender equality in rights, resources and voice. Policy research report no. 21412, The World Bank, Washington, D.C.

Knight, J. and L. Song (2003) Increasing urban wage inequality in China. *Economics of Transition*, **11**:4, 597-619.

Koenker, R. and G. Bassett (1978) Regression Quantiles, *Econometrica*, 46:33-50

Lemieux, T. (2006): The Mincer equation thirty years after Schooling, Experience and Earnings. In: S. Grossbard-Shechtman (Ed.) *Jacob Mincer: A Pioneer of Modern Labour Economics*, Springer Verlag.

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

Machado, J. and J. Mata (2005) Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, **20**:4, 445-465.

Maurer-Fazio, M. (1999) Earnings and education in China's transition to a market economy Survey evidence from 1989 and 1992. *China Economic Review*, **10**:1, 17-40.

Oaxaca, R. (1973) Male-female wage differentials in urban labor markets. *International Economic Review* **14**, 693-709.

Nguyen, B. T., J. Albrecht, S. C. Vroman and D. Westbrook (2007) A quantile regression decomposition of urban-rural inequality in Vietnam. *Journal of Development Economics*, **83**:2, 466-490

Saha, B. and S. Sarkar (1999) Schooling, informal experience, and formal sector earnings: A study of Indian workers. *Review of Development Economics*, **3**(2): 187-199.

Wu, X. and W. Xie (2003) Does the market pay off? Earnings returns to education in urban China. *American Sociological Review*, **68**(3), 425-442.

Zhang, J., Y. Zhao, A. Park and X. Song (2005) Economic returns to schooling in urban China, 1988 to 2001. *Journal of Comparative Economics*, **33**:4, 730-752.