

Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988-2007 (Preliminary)

Xin Meng*

Kailing Shen[†]

Sen Xue[‡]

September 11, 2009

Abstract

In the past 20 years average real earnings of the Chinese urban male workers increased by 350 per cent. Accompanying this unprecedented growth is a considerable increase in earnings inequality. Between 1988 and 2007 the variance of log earnings increased from 0.27 to 0.48, a 78 per cent increase. Using a unique set of repeated cross-section data this paper examines the causes of the increase in earnings inequality. We find that the major changes occurred in the 1990s when the labour market moved from a centrally planned system to a market oriented system. The decomposition exercise conducted in the paper identifies the factor which drives the significant increase in the earnings variance in the 1990s to be the increase in the within education-experience cell residual variances. Such an increase maybe explained mainly by the increase in the price for the unobserved skills. When an economy shift from an administratively determined wage system to a market oriented one, rewards to both observed and unobserved skills increase. The 2000s saw a slowing down of the reward to both the observed and unobserved skills due largely to the college expansion that occurred at the end of the 1990s.

JEL classification numbers: J31

Key Words: Earnings inequality, China

*Economics Program, Research School of Social Sciences, Australian National University, Canberra 0200, Australia. E-mail: Xin.Meng@anu.edu.au Corresponding author.

[†]WISE, Xiamen University, Xiamen, China. E-mail: kailing.shen@gmail.com

[‡]WISE, Xiamen University, Xiamen, China. E-mail: xs225002@126.com

1 Introduction

In the past 20 years China has had unprecedented economic growth. In 1988, China's per capita GDP averaged \$377 and by 2007 this value had increased to \$2604, almost 7 times of that 20 years ago (UNdata). The last 20 years has also seen considerable institutional changes of the economy which switched from a centrally planned to a market oriented system, a large scale rural to urban migration, and a huge educational expansion. In 1988 around 98 per cent of the urban labour force worked either in the state or the collective sectors, only 10 per cent of the urban labour force had a college or above degree, and rural to urban migration was almost non-existent. By 2007, employment in the state or collective sectors dropped by 36 percentage points to 63, college and above degree holders increased by 30 percentage points, while around 140 million rural labour force accounting for one third of the urban labour force moved to cities to work. All these institutional and structural changes should also have significant impact on the distribution of the earnings of urban local workers.

This paper sets to document and investigate the impact of these changes on urban male workers' earnings¹ and their earnings dispersion. Using a unique set of repeated cross-sectional data we show that the last 20 years has seen a quadrupling of the average male annual earnings and a almost doubling of the variance of the log male annual earnings in urban China. We find that there are hardly any changes in earnings inequality in the late 1980s to early 1990s, but then the earnings variance increased from 0.23 in 1992 to 0.47 in 2001. Afterwards, the increase in earnings inequality has been gradual up until 2007.

Part of the significant increase in earnings inequality can be explained by the switching of the urban labour force from the state/collective to the private sectors and part of it is due to the increase in the return to education and the reduction of the return to experience. But the most part of the changes is related to the significant increase in within group residual variances, which has always been referred to as the price for unobserved skills in the literature (see, for

¹“Urban” in China has a special meaning. Since the early 1950s, China implemented a household registration system, whereby individuals born in a city are given an “urban” household registration, otherwise a “rural” household registration. Such a system has not changed for the past 60 years in that individuals who were born to parents with “rural” household registration are still given a “rural” household registration even though he/she might work in cities. In this paper we use the term “urban workers” to refer to workers who have “urban” household registration.

example, Juhn, Murphy, and Pierce, 1993 and Lemieux, 2006).

We explore the detailed changes in institutional settings and supply of skilled and unskilled workers during different period and conclude that in the 1990s the changes in the within group residual variances are probably closely related to the increase in the price for the unobserved skills. The changing trend in the 2000s, however, is mainly driven by the significant increase in the supply of educated workers. The increase in the supply not only suppressed the price for the observed and unobserved skills but also changed the distribution of the the within group unobserved skills. The combination of these two effects may be used to explain the slowing down in the growth of the within group residual variance.

The paper is structure as follows. The next section presents background information, which may affect earnings inequality, such as changes in labour market institutions and supply of and demand for skilled and unskilled labour during this period. Section 3 introduces the data used and describes changes in earnings inequality during the period. Section 4 discusses the methodology. Section 5 investigates the changing pattern of earnings inequality and examines the determinants of these changes. Conclusions are given in Section 6.

2 Background

Under the planned economic regime, urban workers had lifetime employment, whereby their jobs were assigned by the state, firms were not allowed to fire workers, and wage levels were centrally determined. The market oriented economic reform which began in the agriculture sector at the end of 1970s took a long time to spread to the urban sector. By the end of 1980s, although some progress of marketization was made in the goods market, urban labour market reform lagged behind. Neither job mobility nor wage determination had changed much from the planned economic regime except that enterprises had the right to retain a proportion of their profits which were often used to pay bonuses to workers (Korzec, 1992; White, 1988; Takahara, 1992 and Meng, 2000).

The early 1990s saw the speed-up of marketization in the goods market and an increase in dispersion of profitability within the state sector. Studies find that in the early 1990s wage

variations depended heavily on firms' profitability in the state sector (Meng, 2000).

It took the financial collapse of many state owned enterprises to initiate the real changes in the urban labour market. In the mid 1990s, more than 40 percent of the state owned enterprises were making losses. To vitalize the economy, the government finally made it legal for the state enterprises to become bankrupt and to fire unwanted workers. It was not until 1997 that a policy of "keep the 1000 large scale state owned enterprises and let go the small and medium loss making firms" was implemented and in response there was a significant shift of labour force from the state to the non-state sector. In 1997 around 12 million urban state sector workers were laid off (Meng, 1998).

It was against this background that more and more workers shifted to private sector employment. Panel A of Figure 1 presents the share of the state sector employment for the total urban workers and for those who had 5 or less years of work experience. It shows that since 1997-98, the proportion of workers employed in the state sector has been reducing fast, especially for new entrants to the labour market. In 1997 around 75 per cent of the new entrants were employed in the state sector and by 2007 this figure had reduced to 40 per cent.

While most urban people were still working in the state/collective sectors, the private sector employment grew quickly. Since the early 1990s, China has embarked on an export-oriented economic growth strategy. The growth of the FDI and joint venture firms increased demand for unskilled workers in cities. This increase was met by the large scale inflow of rural to urban migrants. According to the data available to us, in 1994 there were around 37 million rural migrant workers in cities, by 2007 this had increased to 135 million (See Panel C of Figure 1). The most significant inflow occurred at around the end of 1990s. The increase in migrant labour supply was accompanied by a high proportion of the urban labour force exiting from the labour market and a shift of urban workers from unskilled production to service and clerk occupation (panel B of Figure 1). Panel C of Figure 1 shows the increase in working aged (16-60) urban non-student population who were not in the labour market, together with the number of rural migrants working in cities for the period 1988 to 2007. The two trends go up almost hand-in-hand.

Another important change in China, which might have a significant impact on urban wage

inequality, is the expansion of higher education. Education levels of urban Chinese have been affected by significant political changes. During the 10 year Cultural Revolution (1966-1976), schools were closed for many years. Universities did not begin normal student recruitment until 1977. As a result, a entire cohort of youth missed schooling for various number of years (Meng and Gregory, 2002 and 2005a). Since then the government has gradually put more emphasis on education. 3 year college and 4 year university admissions increased gradually in the 1990s and in 1998-99 the it was decided to increase 3 year college and 4 year university enrollments by as much as 40 per cent (panel D of Figure 1). The university expansion significantly increased the supply of the skilled workers from the early 2000. Panel E of Figure 1 depicts the change in the share of employed males with different level of education. At the beginning of the period, there was more than 50 per cent of the male workers with junior high school or below education and by the end of the period this ratio had dropped to around 20 per cent. A reversed trend is found for the college or university degree holders. In 1988 a little above 10 per cent of male workers had a college or university degree and by the end of the period the ratio increased to 40 per cent.

Finally, because of the introduction of the One-Child policy at the end of the 1970s, the urban population has aged significantly (panel F of Figure 1). The proportion of young workers (with 1 to 10 years of experience dropped from around 30 per cent of the total workforce at the beginning of the period to just below 20 per cent at the end of the period. The proportion of those with 21 to 30 years of experience increased from 23 per cent to 34 per cent.

We now turn to analyse how these changes affected the earnings distribution in the past 20 years.

3 Data and the general picture of earnings inequality

3.1 Data

The data used in this study come from the Urban Household Income and Expenditure Survey (UHIES) 1988-2007. The survey was conducted by the National Bureau of Statistics annually since 1956 and was resumed in 1980 after its suspension during the Cultural Revolution—1966-

1976 (Fang et al. 2002). The electronic data are available from 1986. The survey samples households with Urban Household Registration for every province in the nation. We use the data from 16 of the 31 provinces including Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Sichuan, Chongqing, Yunan, and Gansu.

This is the only data set in China which goes back 20 years to record annually the earnings changes in the extraordinary growth period. The survey was conducted for the purpose of monitoring income and expenditure changes for households whose household registrations (Hukou) are located in urban areas. It uses a diary record to collect individual earnings, other forms of income, household income and expenditure related data. Basic individual demographic and human capital variables are also available. The data have been widely used to analyse income inequality, poverty, rates of return to education, and household savings in China (see, for example, Han, Wailes, and Cramer, 1995; Fang et al., 2002; Gibson, Huang, and Rozelle, 2003; Meng, Gregory, and Wang, 2005b; Zhang, Zhao, Park, and Song, 2005; Meng, Gregory and Wan, 2007; and Chamon and Prasad, 2008).²

The UHIES questionnaire changed four times in 1988, 1992, 1996, and 2002. The important change for this paper occurred in 1988. In 1986 to 1987, the only income information collected at the individual level is the monthly standard wage. From 1988 onwards, the survey collect individual work related earnings including standard wages, floating wages, bonuses, subsidies, and other work related wages. Because the earnings measured in 1986 and 1987 survey are only subset of the earnings, these years are excluded from the analysis.

Another data related issue worthwhile mentioning is the change in sampling procedure since 2002. Before 2002 the survey only sampled households with urban household registration. Since 2002 the sampling covers all households with a residential address in urban cities where the survey is conducted. However, majority of rural-urban migrants either live in the periphery of cities where rents are cheaper or in dormitories or workplaces such as construction sites. Only a very limited number of households with rural household registration are included in the survey. For example, in 2003 there are around 2 per cent individuals with rural household

²The survey design and sampling proceduel have been discussed in great details in the above mentioned papers.

registration in the sample. To keep the consistency in sampling population, individuals with rural household registration in the 2002-2007 surveys are excluded. In addition to the change in sampling coverage, sample size also more than doubled since 2002.

As the study is on male earnings inequality, the sample includes male individuals aged 20 to 60 who are working and have a positive wage at the survey year.³ Excluding missing values, the final sample size ranges from 7400 to 26,900. Table 1 presents the summary statistics for the data. It shows that over the 20 years, the mean age of the male workers increased slightly from 39 to 42. Similarly, actual and potential work experience⁴ all increased slightly. Further, the average years of schooling increased from 11 years to almost 13 years, while the proportion of workers with 3 year collage and above degree increased from 16 per cent at the beginning of the period to 41 per cent at the end of the period. The most significant changes is the increase in male annual real earnings⁵, which increased from 4,379 yuan in 1988 to 19,520 yuan in 2007, an increase of 346%.

3.2 General picture of earnings inequality

The above discussion focuses mainly on the mean values. As discussed in Section 2, the significant increase in mean earnings over the last 20 years was accompanied by the changes in labour market institutions and the structural change in labour demand and supply. These latter changes brought about the changes in the distribution of the earnings.

Table 2 presents summary measures of earnings inequality for the 20 years, including Gini coefficient, the ratio of the 90th to the 10th percentiles of real annual earnings, and the variance of the log(real annual earnings). The table shows that overall earnings inequality increased significantly by any measure. The Gini coefficient increased from 0.23 to 0.37, an increase of 58 per cent. The ratio of the 90th to the 10th percentiles was 2.9 times in 1988 and increased to 5.4

³In urban China majority of labour force who work works full time. For example, using CHIPS 1995 and 2002 data where days and hours worked variables are available, we found that in 1995 86% of working population in urban China worked 7 to 8 hours and 97% worked 7 or more hours a day. 93% worked 5 to 6 days and 99 per cent worked 5 or more days a week. In 2002 the proportion worked 7 or 8 hours a day is 80%, 7 or more hours a day is 95%, whereas the proportion worked 5 to 6 days is 87% and 5 or more days is 97%.

⁴Potential work experience is calculated as: Age-Years of schooling-7.

⁵Real earnings are calculated using nominal earnings deflated by provincial level CPI which set the 1995 as 100.

times in 2007, an increase of 87 per cent. Similarly, the variance of the log real annual earnings increased by 76 per cent over 20 years. Such an increase is much higher than that found for the US. Over the period of 26 years between 1963 to 1989, there is a 72 per cent increase in variance of log weekly earnings (Juhn, Murphy, and Pierce, 1993).

Figure 2 shows the earnings inequality from different aspects. All of them convey a same story of a fast increasing in earnings inequality. Below we pay particular attention to panels C, D and E.

Panel C presents the change in real earnings for the 10th, 50th, and the 90th percentiles. We observe that in the first few years the earnings of the three groups went up almost equally. Between the early to the end of 1990s the 90th percentile went up much more than the 50th and the 10th percentile groups. Over this period, the bottom income group (the 10th percentile) hardly changed from its early 1990s level. This pattern coincide well with the fact that during this period the drastic state enterprise reform led to a significant drop in earnings for the loss making firms (references). After year 2000, the earnings of the unskilled workers started to grow again, but the rate of change was slower than that for the 90th percentile.

Panel D of Figure 2 shows the ratio of the final years and the base years earnings for each percentile to indicate the 20 years percentage changes for the entire distribution. To reduce measurement error problem, we take the ratio of the average of the last three years (2005-2007) and that of the first three years (1988 to 1990). The scatter plot is the actual changes while the solid line shows the linear prediction. The graph indicates that over the 20 years period the real earnings at the bottom more than doubled, while at the top it increased by almost 4 to 5 times.

The final panel (Panel E of Figure 2) illustrates the same ratios for different periods. We see different patterns over time. The late 1980 to early 1990 period had very small changes in earnings inequality. The 1990 decade has the largest increase in inequality. In this period there is a real earnings drop for the very low earnings group but more than 60 per cent increase for the very high earnings group. The late 1990 to the early 2000 also witnessed large increases in inequality but that was mainly driven by the larger increase in earnings at the top of the distribution. During the early to the mid 2000, earnings increased most dramatically relative

to other periods, but the increase is almost equally distributed and earnings inequality did not change much.

The above descriptive analysis focuses mainly on the overall inequality. As significant institutional and structural changes have occurred in the Chinese urban labour market, the changes in earnings inequality may have affected different groups differentially as indicated in Section 2. Figure 3 presents the variance of the log real earnings by level of work experience, education, occupation, and sector of employment. The graphs indicate earnings inequality for within groups. Panel A shows that earnings inequality is more serious among younger workers than among older workers. However, since the late 1990s, there has not been further widening of earnings inequality among younger workers, whereas for the older workers the inequality continued to increase until 2004. This pattern of wider inequality among the less experienced is unusual from the point of view of western labour markets, where within group inequality is usually higher among the experienced (Juhn, Murphy, and Pierce, 1993 and Lemieux, 2006). This difference, perhaps, is the result of the significant labour market transition from a planned to a market oriented economy. It is the young workers who are more likely to enter into the market oriented private sector, where earnings dispersion has been much wider than that in the public sector (see Panel C of Figure 3). Indeed, even at the beginning of the period (1988), 27 per cent of the group with 1-10 years of experience was not working in the state sector, while for the older group the ratio was 10 per cent. Over time this gap increased. By the end of the period 52 per cent of the young group are not working in the state sector, whereas for those with more than 20 years of experience the proportion is 35 per cent. Another reason for the young to have a larger within group earnings variance may be related to the steepness of the age-earnings profile during the first 10 years of their career.

Panel B of Figure 3 depicts the variance of log real earnings for the educated (3 year college and above) and the uneducated groups (junior high school and below). Unlike many western countries, earnings inequality is much higher among the uneducated group. This, perhaps, is related to the fact that during the Cultural Revolution (1966-1976) entire cohorts of youth were not able to go to school and hence were forced into the less educated group even though their innate ability might otherwise allow them to enter the educated group had they been

given the opportunity to do so. Thus, within the uneducated group the ability distribution should be wider than that in the educated group. Between the 1980s and the 1990s, as China moved from a planned to a market oriented economy, ability was gradually being rewarded and hence we observe the sharp increase in within group earnings inequality for both the educated and uneducated during this period. However, as the Cultural Revolution cohorts gradually retire and more educated young people enter the labour market, especially after the 1990s' university expansion, the inequality of earnings within the less educated group begins to drop while within the educated group inequality continues to increase. This changing pattern of the ability composition within the educated and uneducated groups will remain an important issue throughout the paper.

Panel C of Figure 3 investigates the earnings dispersion within the state (state and collective sectors) and non-state sectors. As expected, the earnings dispersion within the non-state sector is much larger than in the state sector. From the mid 1990s onwards, however, the inequality within the non-state sector dropped significantly for reasons unclear to us at this point. Note that the sample size for the non-state sector is very small for the data between 1988 and 1993 and relatively small between 1994 and 1996. Thus, it is possible that the extremely high variance during the period of 1991 to 1996 is due to sample noise and should not be taken seriously. Within the state sector the earnings inequality started off from a fairly low level and increased continuously during the 1990s. Since the early 2000 the within state sector earnings variance has kept constant.

The last panel (Panel D) examines different pattern of inequality for different occupation groups. The service/clerk workers and production workers exhibit larger within group earnings inequality than the highly educated professional and managerial group. The decline of earnings variance among the production workers from 1999 onwards may be related to the large scale inflow of the rural migrant workers and the subsequent increase in the proportion of urban local workers exiting from the labour market.

The general inequality picture presented above indicates that there has been a considerable increase in earnings inequality and the increase occurred mainly in the 1990s. The changes in inequality among different groups of labour force, however, is not the same.

4 Methodology

The increase in earnings inequality is often decomposed into the change in the price of observable characteristics, the change in the distribution of observable characteristics, and the change in the price of the unobservable ability (See, for example, Juhn, Merphy, and Pierce, 1993; DiNardo, Fortin, and Lemieux, 1996, Lemieux, 2002 and 2006). We mainly follow Lemieux (2002) and Lemieux (2006) to decompose the change in variance of earnings. Consider the following earnings equation:

$$y_{it} = x_{it}\beta_t + \epsilon_{it}, \quad (1)$$

where y_{it} is the log of the annual earnings of individual i at year t ; x_{it} is a vector of observable characteristics (education and labour market experience for the baseline model and education, experience, and state sector employment for the extended model); β_t is a vector of return to observable characteristics; and ϵ_{it} is the error term. Lemieux (2002) shows that the variance of the log earnings can be written as:

$$V = Var(y_{it}) = \sum_i \omega_{jt} (x_{it}\widehat{\beta}_t - \bar{x}_t\widehat{\beta}_t)^2 + \sum_i \omega_{it} u_{it}^2 = \sum_j \theta_{jt} (\bar{y}_{jt} - \bar{y}_t)^2 + \sum_j \theta_{jt} \sigma_{jt}^2, \quad (2)$$

where ω is the sample weight, $\widehat{\beta}$ is the vector of ols estimates of β , and u is the residual from the ols regression of equation (1). In the second part of the equation (2), the subscript j indicates the number of skill cells in the sample (for example, if x is a vector of experience and education categories, j will be the number of experience-education cells in the sample); σ_{jt}^2 is the within cell residual variance, and θ_j is the proportion of observations in cell j of the sample (in the case where x are all dummy variables, θ will be \bar{x}).

To decompose the change in the variance of earnings between year t and year s , Lemieux (2002) proposes two counterfactual variances. The first is to replace $\widehat{\beta}_t$ in Equation (2) with $\widehat{\beta}_s$ which is obtained by estimating Equation (1) with period s data. Thus, the first counterfactual variance, denoted as V^a , is defined as:

$$V^a = \sum_i \omega_{jt} (x_{it} \hat{\beta}_s - \bar{x}_t \hat{\beta}_s)^2 + \sum_i \omega_{it} u_{it}^2 = \sum_j \theta_{jt} (\bar{y}_{js} - \bar{y}_t^a)^2 + \sum_j \theta_{jt} \sigma_{jt}^2. \quad (3)$$

The second counterfactual variance, denoted as V^b , is to replace θ_{jt} in Equation (3) by θ_{js} , which is the actual proportion of observation in cell j for period s :

$$V^b = \sum_i \omega_{js} (x_{is} \hat{\beta}_s - \bar{x}_s \hat{\beta}_s)^2 + \sum_i \omega_{is} u_{it}^2 = \sum_j \theta_{js} (\bar{y}_{js} - \bar{y}_s)^2 + \sum_j \theta_{js} \sigma_{jt}^2 \quad (4)$$

Using the two counterfactual variances, V^a and V^b , we can decompose the change in variance of earnings between periods t and s into three components: 1. a portion which is due to the change in the return to the observables ($\Delta \hat{\beta}$); 2. a portion which is due to the change in the composition of the sample ($\Delta \theta_j$); and 3. a portion which is due to the change in the within cell residual variance ($\Delta \sigma_j$). Thus, $V_t - V_s$ may be written as:

$$V_t - V_s = \underbrace{(V_t - V_t^a)}_{\text{Observable price}} + \underbrace{(V_t^a - V_t^b)}_{\text{Composition}} + \underbrace{(V_t^b - V_s)}_{\text{Within cell variance}} \quad (5)$$

Lemieux (2006) further investigates the residual variance component and writes it as follows:

$$\text{Var}(u_{it}) = p_t^2 \sum_j \theta_{jt} \sigma_{jt}^2 \quad (6)$$

where p_t is price for unobservables. He then assumes that the distribution of unobserved skills among workers within each skill cell j is stable over time, thus, $\sigma_{jt}^2 = \sigma_{js}^2 = \sigma_j^2$. Because of this, Equation (6) can be re-written as $\text{Var}(u_{it}) = \sum_j \theta_{jt} V_{jt}$, where V_{jt} varies only if the price for unobservables vary over time.

The assumption that the distribution of the unobservable skills is constant over time is unrealistic for China. One example may be worth mentioning. As discussed in Section 2, the significant expansion of the college enrollment in 1999 enabled a large number of less capable people to obtain degrees. Thus, within the college cell the unobserved skill distribution should widen significantly, and, at the same time, the distribution of the unobserved skill distribution for the low skilled group (junior high and below education group) should narrow. Hence, the

last term in Equation (5) does not only represent the change in the price for the unobservables but also the change in the distribution of unobservables.

5 Earnings inequality

5.1 Results from the earnings equations

We first estimate the earnings equation (Equation (1)) for each of the 20 years between 1988 and 2007. To reduce measurement error problems, we exclude the top and bottom 1 per cent of the sample. Two specifications of the earnings equation are used. The baseline model controls for four categories of education (3 year college and above, technical school, Senior high school, and junior high school and below), 4 categories of years of work experience (1-10 years, 11-20 years, 21-30 years, and 31 and above years) and provincial fixed effects. This specification is consistent with those used in most studies on the US wage inequality literature. The inclusion of provincial dummy variables is necessary to control for significant spatial price variations (Brandt and Holz, 2006 and Gong and Meng, 2008).⁶ In addition, we also extend the baseline model with an additional dummy variable indicating whether an individual is working in the state sector or not because the switch between the state and non-state sector is probably an important story in the change in earnings inequality.

Tables 3 and 4 present the results for the Baseline and the extended models, respectively. In both cases the omitted category is 1-10 years experience and junior high school and below. The results from Table 3 shows that there seems to be a very significant returns to years of experience. The longer the experience, the higher the return throughout the 20 year period. However, relative to the group with 1-10 years of experience, the change in the return to experience varies quite a lot for individuals in different experience groups. In the late 1980s and early 1990s those with 11-20 years of experience earned around 35 per cent more than those with 1-10 years of experience. Since the early 1990s the return to experience for this group has reduced to around 25-28 per cent. For those with 21-30 and above 30 years of experience,

⁶We assume that provincial variations are fully captured by the changes in β s and Xs and will have no effect either on residual variance distribution across groups nor within groups.

the reduction in relative returns has been more dramatic and continuous. For individuals with 21-30 years of experience the returns reduced from 48 per cent in 1988 to 34 per cent in 2007, while for those with above 30 years of experience the change is from 55 per cent to 35 per cent. This phenomenon is probably related to the fact that during the planned economic era, wages were administratively determined and seniority was one of the most important consideration in the earnings determination process. This is less true in a more market oriented economy (Meng and Kidd, 1997 and Meng, 2000).

While the return to experience has been reducing, the return to education has been increasing. Compare to the junior high school and below group, the Senior high, technical school, and 3 year college and above groups in 1988 earned 2, 7, and 13 per cent more. By 2007, the same ratios increased to 17, 27, and 56 per cent, respectively.

Panels A and B of Figure 4 present these changes visually. The most important message from Panel A is the continuous decline in the rate of returns to experience, especially for the 21 and above years of work experience groups. Panel B shows that the most significant increase in returns to college and above degree occurred during the 1990s with a couple of years slowing down in the mid 1990s. By the early 2000 the trend has flattened out. This, perhaps, is related closely to the college expansion. The patterns for the technical school and senior high school are similar but at a lesser extent.

The extended model presents similar picture of the change in rates of return to education and work experience. The additional variable, the employment in the state or collectively owned enterprises, has had dramatic changes in its payoffs. During the late 1980s and early 1990s, employment in the state or collective sector had a positive earnings payoff. This payoffs had disappeared in the 1990s and in the mid 1990s it became negative as a large proportion of these firms were making losses. It was not until after the early 2000s that the payoff to the state and collective employment returned to positive again (see Panel C of Figure 4).

Another way to gauge the contribution of each set of the variables included in the regressions in explaining the total earnings variation is to derive a bounded range of the contributions using partial and marginal R^2 s for each set of them (see Dickens and Katz, 1987). We plot out the

upper bound (the partial R^2 s) of the contributions.⁷ These are presented in Panel D of Figure 4. Some striking features are worth noting. First, in the late 1980s experience alone explained around 25 per cent of the earnings variances and the effect reduced dramatically over the time and by 2007 it only explained less than 3 per cent of the earnings variation, suggesting a disappearing of a seniority driven wage determination system. Second, in the 1990s provincial dummy variables explained around 20 per cent of the earnings variations and it was not until the late 1990s this effect reduced and stayed at around 10 per cent of the earnings variation for the rest of the period. Finally, education did not become an important story until after the late 1990s.

Another issue worth mentioning is the extent to which earnings variations can be explained by all the observed characteristics. The estimated earnings regression results show a decline of the adjusted R^2 s. In the late 1980s and early 1990s the adjusted R^2 s are around 0.33 to 0.38. From the mid 1990s the adjusted R^2 s starts to reduce and by the end of the period, the observable characteristics can only explain 23 to 25 percent of the earnings variations. Thus, over time, the residual component of the earnings variations has increased significantly.

5.2 Decomposing earnings inequality

Before engaging in variance decomposition, we first plot the change of the total variance, the variance of the portion which can be explained by the changes in characteristics and in the rate of return to characteristics, and the residual variance. Panel A of Figure 5 presents the traditional Blinder style decomposition of the total earnings variance. It shows that the total variance has increased from just below 0.2 in 1988 to almost 0.4 in 2007. Of this change a significant portion comes from the increase in the residual variance. We also explore the change in variance of the explained portion (Panel B of Figure 5) and find that most of the explained variance is due to changes in the rate of return to characteristics.

Next we apply Lemieux's (2002) decomposition method to understand what generated the large increase in earnings inequality, in particular, residual inequality, over the last 20 years.

⁷We also calculated the lower bound of the impact of each set of variables on earnings variation by adding each set of the variables on the existing sets of variables to see its marginal contribution to the R^2 s. These results are similar to those presented in Panel D of Figure 4 and are available upon request from the authors.

Table 5 presents the change in variance of earnings between 1988 and 2007, as well as four different time periods within this 20 years, that is between 1988 and 1993, 1993 and 1998, 1998 and 2003, and 2003 and 2007. The top panel of the table uses the basic model and 1988 as the base year. The first column presents the change in variance between 1988 and 2007. During the 20 year period, the total variance of log earnings increased by 0.21, which is a 76 per cent increase to its 1988's level of 0.27. Of the total change in variance, 91 per cent is due to the within group residual variance increase. The composition effect, that is the change in the proportions of people in different education-experience cells, contributes to a narrowing of the earnings inequality by 13 per cent. The change in returns to observable characteristics contributes to the increase in the earnings variation. These two effects, however, almost cancel each other out.

The next four columns split the period into four sub-periods. In the 1988-1993 period urban economy operated under the old planned economy system and earnings inequality increased very little. Of the very small change, all three components (changes in the price of the observables, composition effect, and the changes in the within group variances) account for almost equal amounts, with the composition effect contributing to the narrowing while the other effects contribute to the enlarging of the earnings variation.

The urban economic reform took place mainly in the 1990s, when the reward to observable skills increased the most significantly. During this period many state owned factories were making losses and more able workers switched to the private sector while less able workers exited from the labour market. Accompanying this process was the large scale rural-urban migration which also contributed to changes the urban labour market. It was during this period that the earnings inequality widened most significantly. Over the 6 years between 1993 and 1998, the variance of the log earnings increased from 0.21 to 0.32, which accounted for more than half of the total increase in the variance of log earnings over the entire 20 years. Of the total change, 95 per cent can be attributed to the increase in within group residual variances.

The 1998-2003 period continued the trend of large inflow of rural migrants to the urban labour market. The variance of the log earnings increased by 0.08, which is the second largest change among the four sub-periods. Once again the increase in the within group residual

variances explains more than 100 per cent of the change in the earnings inequality.

The 2003-2007 period the earnings inequality declined marginally and for the first time the change in the within group residual variances contributed to a decline of earnings inequality. This, may, to a large extent, be related to the significant inflow of college and above graduates to the labour market from 2002. The increase in the supply of skilled labour should have played an important role in the slowing down of the trend towards a widening in earnings inequality. We will come back to this point later in this section.

These results suggest that over the last 20 years the majority of the increase in earnings inequality is due to the increase in within group residual inequality. Panel A of Figure 6 presents the decomposition results for each of the 20 years, where 1988 is used as the base year. Thus, the changes in the return to observable characteristics are predicted using every year's regression coefficients but 1988's characteristics. Similarly, the changes in composition effects are calculated relative to 1988, while the changes in within group residual variances are also calculated as each year's within group variances relative to those of 1988. The trend presented in this figure confirms the dominant role the changes in the within group variance played in explaining the increase in earnings inequality during this period, especially during the 1990s. By the early 2000 the continuous increase in within group residual variances has stopped.

The second panel of Table 5 presents the same decomposition results using 2007 as the base year. As all decompositions suffer from the index number problem, this panel is presented so to allow us to gauge the bound of the effects. The trends indicated here (and in Panel B of Figure 6) are largely consistent with those obtained from using the 1988 as the base year, except for the last period. Luckily this is the period which the earnings variation hardly changed.

The last panel of Table 5 shows the results from the model including an additional dummy variable indicating whether an individual is employed in the state or collective sector. The decomposition uses 2007 as the base year.⁸ Relative to the results presented in the second panel of this table, we find that including the state/collective employment increases the composition effect significantly. When not controlling for the state/collective employment, the composition

⁸In 1988 98 per cent of the sample were employed either in the state or collective sectors. Thus, if we use 1988 as the base year many education-experience-sector of employment cells will be empty, which makes the decomposition impossible.

effect contributes to a 4 per cent reduction of the change in earnings variations. Once the state/collective employment is controlled for, the composition effect becomes positive 23 per cent, suggesting that the shift from the state to non-state employment contributes to 27 per cent (4 per cent plus 23 per cent) of the total increase in the earnings variations. This is quite a significant effect. We also find that the major changes occurred between 1998 and 2003, which coincide with the significant reduction in the state/collective employment as presented in Panel A of Figure 1 in Section 2.

Even though including state/collective employment increases the contribution of the composition effect, the general dominance of the within group residual variance over the period does not change. For the entire period it still accounted for 65 per cent of the changes in the variance of the earnings and for the period 1993 to 1998 it accounts for 87 per cent of the total change. This can also be seen from Panel C of Figure 6 which compares the decomposition results from the baseline model using 2007 as the base year with that of using the extended model and 2007 as the base year.

5.3 What caused the increase in within group residual variance?

The question naturally arises as to what is the driving force for the significant increases in within group residual variances? In the literature changes in within group residual variances are fully attributed to changes in price for the unobserved skills. Juhn, Murphy and Pierce (1993) and Lemieux (2002) both assumed that the within group distribution of unobserved skill is held constant over time. Given China's special institutional changes over the past 20 years we believe that this assumption is unrealistic.

One possible cause for over time changes in within group unobserved ability distribution is the significant changes in education attainment. As discussed in the background section, the education expansion occurred gradually during the mid to late 1990s and in 1998-1999 there was a sudden increase in college and university enrollment. In 1988 only 16 per cent of the total male labour force had college and above degree and this ratio increased to above 40 per cent in 2007. While the proportion of technical school and senior high school graduates did not change much over the same period, the ratio of the work force with junior high school or below

education reduced from over 48 per cent in 1988 to less than 24 per cent in 2007.

How should the increase in the college enrollment affect the within group distribution of the unobserved skills? The following example may illustrate this clearly. Assume that unobserved skill distributions for two periods for the whole population are constant and that the total population is divided into two groups: the educated and uneducated. Assume that in period one only 10 per cent of the population is in the educated group and in period two this increases to 40 per cent. In this situation, the within group distribution of unobserved skills for the educated group should be wider in the second period than in the first period, while it should be narrower for the uneducated group in the second period than that in the first period. This is because only the relatively able people can pass the college entry exam and receive college and above education. The change in the distribution of unobserved skills may be shown in Figure 7, where the normal distribution curve represents the distribution of unobserved skills for the entire population for both period 1 (t) and period 2 ($t + 1$). The two gray bars indicate the division between the educated and uneducated groups in period 1 ($high_t$) and period 2 ($high_{t+1}$). As the society decides to educate more people it moves the “standard” for being in the educated group lower (the gray bar moves to the left). As a result, the distribution of unobserved skill within the educated group increases from the distance “educated t ” to the distance “educated $t+1$ ”, while that of the uneducated group decreases by the the same amount (from “uneducated t ” to “uneducated $t+1$ ”).

This simple example shows that when the proportion of college degree holders in the total labour force doubled while proportion of labour force with junior high school education reduced by half it is not possible to assume that the within group distribution of unobserved ability is held constant and all the changes in the within group residual variances are due to the change in the price for the unobserved ability. Based on the timing of the education expansion, we would expect that in the 2000s the change in the within group distribution of unobserved skills to play an increasing role in explaining the change in the within group residual variance.

With this in mind, we plot the changes in residual earnings variances by education and experience cells (Figure 8). While Panels A and B of Figure 8 show the different trends in the residual variance for different experience groups within the college and above education and the

junior high and below education groups, respectively, Panel C and D present the change for different education levels within the 1 to 10 years of experience and 21 to 30 years of experience groups.

The top two graphs reveal the following information. First, they indicate that within either the most or the least educated group, the young group (with 1-10 years of experience) exhibits the highest within group residual variances. This is a continuation of the pattern observed in the total within group variances, suggesting that the return to observable characteristics and the composition differences between the young and old groups cannot explain the large gap in variance between the young and the old. Indeed, the partial adjusted R²s for work experience have been reducing dramatically over time (see Panel D of Figure 4). However, why the residual inequality is higher for the young than for the old remains a puzzle to us.

Second, there is a significant increase in the within group residual variances during the 1990s and this trend is common for both young and old, and educated and uneducated groups. Given that during this period the gradual college expansion has not yet or just started, it is unlikely that there are significant increases in unobserved skill distributions within each group. There are two possible causes for the increase in the price for unobserved skills. One is the increase in demand for unobserved skills and the other is the changes in the wage determination system due to the shift from a planned to a market economy. The effect of demand change may deserve a separate study. However, based on both the survey data and the aggregated data from the statistical yearbook (NBS, various years) on sector of employment, it was not until 1998 the share of workers in manufacturing sector significantly dropped (see Figure 10). Whereas the major increase in the within group residual variances occurred between 1991 and 1998. Further, the 1990s experienced the most significant labour market reform and earnings changed from administratively determined to a market determined system. Indeed, we observed a considerable drop in rate of returns to experience (seniority) and increase in rate of return to education (Figure 4) during this period, reflecting the switching regims of the wage determination. We, therefore, tend to believe that the significant increase in within group residual variances during this period is closely related to economic reform led increase in returns to unobserved skills.

Third, the 2000s has seen some differential trends. In both educated and uneducated groups

the slowing down in the increase in the residual variances is obvious. Within the college group the between young and old group variance did not change much over the entire period, but we do observe a narrowing of the gap between the residual variance of earnings for the young and for the old for the unskilled group (those with junior high and below education).

Examining panels C and D of Figure 8 reveals a more interesting pattern. For the group with 1-10 years of experience, the difference in variance of earnings between the college and above and the junior high and below narrowed significantly in the 2000s. This pattern, however, is not observed for the old group (21 to 30 years of experience) where in the 2000s the variance between the more and less educated groups has widened. To show this pattern more clearly, we reproduce the two graphs from panels C and D, excluding the technical and senior high school groups (see Figure 9).

These differential trends revealed in Figure 9 allow us to identify the causes of the changes in the within group residual variances for the 2000s. Recall that the college expansion took effect from the early 2000s and that this mainly affected the young group if we assume that the substitution between the young and the old groups is low. The simple example given in Figure 7 suggests that for the young group we should observe a widening of the unobserved skill distribution for the educated group (college and above) and a narrowing of the distribution for the uneducated (junior high and below). This pattern, however, should not be observed for the old group. What we find in Figure 9 is almost exactly this pattern. To show this more clearly we take difference-in-differences of the within group residual variance between the educated and uneducated in the young group and the educated and uneducated in the old group (Figure 11). If the difference between the educated and uneducated in the young group mimic that in the old group we should see the diff-in-diff plot hovering around the zero line. Figure 10 shows that this is the case for the late 80s and most of the 90s, but not for the 2000s. In the 2000s, the difference of within group residual variances between the educated and uneducated for the young group is widened significantly relative to that for the old group.

Although the pattern of changes in within group variances among different groups presented in Figure 9 fits in with our predictions, there are some details which do not fit exactly. In particular, although the left panel of Figure 9 indicates a reduction in within group residual

variance for the uneducated in the young group in the 2000s, it does not show a similar degree of increase in the within group variance for the educated group as we have predicted. Our conjecture for this inconsistency is that, perhaps, the significant increase in the supply of college degree holders have not only suppressed the price for the observable but also unobservable skills. This may be confirmed by the fact that without significant changes in the distribution of the unobserved skills for the educated in the old group, the within group residual variance for this group has also declined. Thus, the decrease in the price for unobserved skills offsets the effect of the increase in the distribution of the unobserved skills for the young educated group. Had there not been a slowing down of the rewards to the observed and unobserved skills, we would have observed a sharp increase in within group residual variances for the educated young group.

6 Conclusions

In this paper we studied the change in male earnings inequality in urban China during a period where China experienced unprecedented economic growth and significant institutional changes. We found that over the 20 years period, earnings inequality increased by almost 60 per cent if inequality is measured by Gini coefficient and 76 per cent if it is measured as variance of log real earnings. Such a significant increase had not been seen in many other countries.

Our analysis indicated that while the changes in prices for observed skills and the changes in composition of male labour force explained some of the changes, more than 80 per cent of the increases in earnings inequality during this period was due to the increase in the unobserved within group residual inequality.

Our analysis suggested that during the 1990s the increases in within group residual variances were mainly due to the increases in returns to the unobserved skills. As the economy transformed from a system where wages were administratively determined to a market system, the reward to both observed and unobserved skills increased.

In the 2000s, however, the trend in within group residual variances changed significantly. This was a result of a combination of the changes in the distribution of the within group unobserved skills and a reduction of the reward to the unobserved (as well as observed) skills

due to the over supply of the labour force with college and above degrees. As a result of the significant expansion of the college enrollment from the late 1990s, the within group unobserved skill distribution widened in the educated group and narrowed in the uneducated group. We found that this change in the distribution of within group unobserved skill could, perhaps, explain a large part of the slowing down of the changes in within group residual variances.

References

- Brandt, Loren and Holz, Carsten A. (2006). "Spatial Price Differences in China: Estimates and Implications." *Economic Development and Cultural Change*, Vol. 55 (1), pp. 43-86
- Chamon, M. and Prasad, E., 2008, "Why are saving rates of urban households in China rising?", *NBER Working Paper Series* No. 14546.
- Dickens, W. T. and Katz, L. F., 1987, "Inter-Industry Wage Differences and Industry Characteristics" in K. Lang and J. S. Leonard (eds) *Unemployment and the Structure of Labor Markets*, Oxford: Basil Blackwell Inc., pp.48-89.
- DiNardo, J., Fortin, N.M., and Lemieux, T., 1996, "Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach", *Econometrica*, 64(5), pp.1001-1044.
- Han, T., Wailes, E.J., and Cramer, G.L., 1995, "Rural and urban data collection in the People's Republic of China", in *the China Market Data and Information Systems, Proceedings of WCC-101 Symposium*, Washington, D.C.
- Fang, C., Zhang, X., and Fan, S., 2002, "Emergence of urban poverty and inequality in China: evidence from household survey", *China economic Review*, 13, pp.430-443.
- Gibson, J., Huang, J., and Rozelle, S., 2003, "Improving estimates of inequality and poverty from urban China's household income and expenditure survey", *Review of Income and Wealth*, 49(1), pp.53-68.
- Gong, C.H. and Meng, X., 2008, "Regional price differences in urban China 1986-2001: Estimation and implication", *IZA Discussion Paper* No. 3621.
- Juhn, C., Murphy, K.M., and Pierce, B., 1993, "Wage inequality and the rise of returns to skill", *Journal of Political Economy*, 101(3), pp.410-442.
- Korzec, Michael, 1992, *Labor and the Failure of Reform in China*, London: MacMillan Press LTD, 1992.
- Lemieux, T., 2002, "Decomposing changes in wage distributions: a unified approach", *Canadian Journal of Economics*, 35(4), pp.646-688.
- Lemieux, T., 2006, "Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill?", *American Economic Review*, 96(3), pp.461-498.

- Meng, X. and Kidd, M., 1997 “Wage Determination in China’s State Sector in the 1980s”, *Journal of Comparative Economics*, 25(3), pp.403-421.
- Meng, X., 2000, *Labour Market Reform in China*, Cambridge: Cambridge University Press.
- Meng, X. and Gregory, R. G., 2002, “The impact of interrupted education on subsequent educational attainment: A cost of the Chinese Cultural Revolution”, *Economic Development and Cultural Change*, 50(4), pp.935-959.
- Meng, X., Gregory, R.G. and Wang, Y., 2005, “Poverty, inequality, and growth in urban China, 1986-2000”, *Journal of Comparative Economics*, 33(4), pp 710-729.
- Meng, X. and Gregory, R. G., 2005, “The impact of interrupted education on earnings: the educational cost of the Chinese Cultural Revolution” *IZA Discussion Papers*, DP No. 2548.
- Meng, X., Gregory, R. and Wan, G., 2007, “China urban poverty and its contributing factors, 1986-2000” *Review of Income and Wealth*, 53(1), pp.167-189.
- Takahara, Akio, 1992, *The Politics of Wage Policy in Post-Revolutionary China*, London: MacMillan Press LTD.
- White, Gordon, 1988. “State and Market in China’s Labor Reforms”, *Journal of Development Studies*, 24, 4:180-202.
- Zhang, J., Zhao, Y., Park, A., and Song, X., 2005, “Economic returns to schooling in urban China, 1988 to 2001”, *Journal of Comparative Economics*, 33, pp.730-752.

Table 1: Summary Statistics, 1988-2007

| | <u>1988</u> | | <u>1989</u> | | <u>1990</u> | | <u>1991</u> | | <u>1992</u> | | <u>1993</u> | | <u>1994</u> | | <u>1995</u> | | <u>1996</u> | | <u>1997</u> | |
|-------------------------------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|
| | Mean | SD |
| age | 39.03 | 10.41 | 39.23 | 10.31 | 39.44 | 10.34 | 38.96 | 9.99 | 39.13 | 10.03 | 39.35 | 9.97 | 39.16 | 9.85 | 39.51 | 9.69 | 39.73 | 9.59 | 39.82 | 9.37 |
| Years of schooling | 10.95 | 2.89 | 11.11 | 2.92 | 11.23 | 2.87 | 11.39 | 2.92 | 11.52 | 2.73 | 11.58 | 2.67 | 11.72 | 2.68 | 11.77 | 2.65 | 11.78 | 2.62 | 11.77 | 2.63 |
| % with 3 year college & above | 0.16 | | 0.18 | | 0.19 | | 0.21 | | 0.24 | | 0.24 | | 0.26 | | 0.26 | | 0.26 | | 0.26 | |
| Actual work experience | 19.40 | 10.44 | 19.96 | 10.35 | 20.29 | 10.38 | 19.85 | 10.09 | 19.95 | 10.15 | 20.20 | 10.04 | 20.11 | 10.00 | 20.46 | 9.82 | 20.68 | 9.78 | 20.71 | 9.65 |
| Potential work experience | 21.03 | 10.92 | 21.07 | 10.90 | 21.19 | 10.88 | 20.56 | 10.63 | 20.59 | 10.57 | 20.76 | 10.47 | 20.44 | 10.40 | 20.74 | 10.22 | 20.94 | 10.19 | 21.05 | 10.04 |
| log real annual earnings | 8.27 | 0.52 | 8.25 | 0.50 | 8.32 | 0.48 | 8.37 | 0.47 | 8.51 | 0.48 | 8.56 | 0.54 | 8.63 | 0.61 | 8.67 | 0.60 | 8.69 | 0.61 | 8.72 | 0.64 |
| Real annual earnings | 4379 | 1978 | 4271 | 1985 | 4537 | 2117 | 4765 | 2369 | 5520 | 3108 | 5991 | 3647 | 6632 | 4349 | 6849 | 4334 | 7052 | 4821 | 7390 | 5150 |
| No. of obs. | 7,890 | | 7,087 | | 7,470 | | 7,587 | | 9,845 | | 9,477 | | 9,443 | | 9,357 | | 9,414 | | 9,376 | |
| | <u>1998</u> | | <u>1999</u> | | <u>2000</u> | | <u>2001</u> | | <u>2002</u> | | <u>2003</u> | | <u>2004</u> | | <u>2005</u> | | <u>2006</u> | | <u>2007</u> | |
| | Mean | SD |
| age | 40.09 | 9.26 | 40.42 | 9.12 | 40.31 | 9.30 | 40.98 | 9.13 | 41.55 | 9.00 | 41.81 | 8.93 | 42.19 | 8.93 | 42.01 | 9.08 | 42.38 | 9.01 | 42.36 | 9.16 |
| Years of schooling | 11.88 | 2.60 | 12.00 | 2.58 | 12.04 | 2.67 | 12.01 | 2.63 | 12.35 | 2.67 | 12.42 | 2.65 | 12.48 | 2.67 | 12.53 | 2.71 | 12.64 | 2.69 | 12.75 | 2.68 |
| % with 3 year college & above | 0.28 | | 0.29 | | 0.31 | | 0.30 | | 0.36 | 0.48 | 0.37 | 0.48 | 0.38 | 0.48 | 0.39 | 0.49 | 0.40 | 0.49 | 0.42 | 0.49 |
| Actual work experience | 20.96 | 9.60 | 21.32 | 9.53 | 20.82 | 9.89 | 21.42 | 9.83 | 21.06 | 9.71 | 21.31 | 9.72 | 21.60 | 9.79 | 21.24 | 9.97 | 21.56 | 9.94 | 21.96 | 10.18 |
| Potential work experience | 21.20 | 9.95 | 21.43 | 9.83 | 21.27 | 10.09 | 21.97 | 9.94 | 22.19 | 9.93 | 22.39 | 9.85 | 22.70 | 9.86 | 22.47 | 10.01 | 22.73 | 9.94 | 22.60 | 10.13 |
| log real annual earnings | 8.76 | 0.66 | 8.84 | 0.66 | 8.90 | 0.70 | 8.98 | 0.69 | 9.14 | 0.66 | 9.20 | 0.69 | 9.31 | 0.68 | 9.40 | 0.69 | 9.49 | 0.69 | 9.64 | 0.69 |
| Real annual earnings | 7725 | 5611 | 8318 | 5329 | 9161 | 6768 | 9875 | 7137 | 11320 | 7932 | 12298 | 9203 | 13740 | 10769 | 15196 | 12098 | 16723 | 13473 | 19520 | 18132 |
| No. of obs. | 9,186 | | 9,043 | | 8,841 | | 8,668 | | 21,581 | | 24,411 | | 25,785 | | 26,701 | | 26,967 | | 24,736 | |

Table 2: Summary measures of earnings inequality, 1988-2007

| | Gini(real annual earnings) | Ratio of 90th to 10th percentile of real annual earnings | Variance(log real annual earnings) |
|------|----------------------------|--|------------------------------------|
| 1988 | 0.232 | 2.879 | 0.275 |
| 1989 | 0.238 | 2.932 | 0.247 |
| 1990 | 0.231 | 2.899 | 0.230 |
| 1991 | 0.234 | 2.739 | 0.222 |
| 1992 | 0.242 | 2.775 | 0.233 |
| 1993 | 0.275 | 3.246 | 0.297 |
| 1994 | 0.309 | 3.917 | 0.367 |
| 1995 | 0.300 | 3.745 | 0.356 |
| 1996 | 0.311 | 3.853 | 0.374 |
| 1997 | 0.320 | 4.011 | 0.413 |
| 1998 | 0.327 | 4.216 | 0.430 |
| 1999 | 0.316 | 4.292 | 0.436 |
| 2000 | 0.341 | 4.881 | 0.490 |
| 2001 | 0.340 | 4.769 | 0.472 |
| 2002 | 0.323 | 4.539 | 0.420 |
| 2003 | 0.354 | 5.362 | 0.541 |
| 2004 | 0.362 | 5.238 | 0.506 |
| 2005 | 0.369 | 5.530 | 0.524 |
| 2006 | 0.367 | 5.493 | 0.502 |
| 2007 | 0.368 | 5.384 | 0.483 |

Table 3: Selected regression results from the baseline model, 1988-2007

| | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
|------------------------|-----------------------|-----------------------|----------------------|------------------------|------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Experience 11-20 years | 0.364*** [0.0108] | 0.346*** [0.0119] | 0.361*** [0.0109] | 0.286*** [0.0107] | 0.267*** [0.00949] | 0.287*** [0.0114] | 0.270*** [0.0128] | 0.278*** [0.0131] | 0.279*** [0.0136] | 0.264*** [0.0146] |
| Experience 21-30 years | 0.484*** [0.0117] | 0.467*** [0.0124] | 0.493*** [0.0110] | 0.422*** [0.0108] | 0.402*** [0.00947] | 0.436*** [0.0112] | 0.439*** [0.0130] | 0.415*** [0.0132] | 0.404*** [0.0135] | 0.390*** [0.0143] |
| Experience >30 years | 0.548*** [0.0130] | 0.539*** [0.0135] | 0.577*** [0.0121] | 0.494*** [0.0124] | 0.475*** [0.0110] | 0.496*** [0.0131] | 0.507*** [0.0152] | 0.484*** [0.0153] | 0.490*** [0.0157] | 0.452*** [0.0170] |
| College | 0.127*** [0.0115] | 0.146*** [0.0118] | 0.164*** [0.0104] | 0.172*** [0.00997] | 0.196*** [0.00860] | 0.239*** [0.0101] | 0.306*** [0.0116] | 0.253*** [0.0116] | 0.269*** [0.0119] | 0.256*** [0.0126] |
| Technical school | 0.0683*** [0.0130] | 0.0821*** [0.0139] | 0.122*** [0.0120] | 0.125*** [0.0116] | 0.128*** [0.0103] | 0.149*** [0.0121] | 0.192*** [0.0141] | 0.164*** [0.0143] | 0.170*** [0.0146] | 0.152*** [0.0159] |
| Senior high school | 0.0210** [0.0103] | 0.016 [0.0108] | 0.0146 [0.00955] | 0.0413*** [0.00945] | 0.0470*** [0.00856] | 0.0698*** [0.0101] | 0.0889*** [0.0119] | 0.0678*** [0.0117] | 0.0834*** [0.0119] | 0.0895*** [0.0126] |
| Provincial dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7729 | 6941 | 7317 | 7435 | 9648 | 9288 | 9254 | 9170 | 9220 | 9184 |
| R-squared | 0.35 | 0.33 | 0.39 | 0.38 | 0.38 | 0.37 | 0.36 | 0.33 | 0.33 | 0.32 |
| | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| Experience 11-20 years | 0.275*** [0.0155] | 0.251*** [0.0162] | 0.278*** [0.0171] | 0.278*** [0.0181] | 0.254*** [0.0110] | 0.292*** [0.0113] | 0.308*** [0.0108] | 0.268*** [0.0106] | 0.288*** [0.0107] | 0.281*** [0.0111] |
| Experience 21-30 years | 0.397*** [0.0151] | 0.362*** [0.0158] | 0.386*** [0.0167] | 0.383*** [0.0175] | 0.328*** [0.0107] | 0.352*** [0.0109] | 0.343*** [0.0105] | 0.317*** [0.0104] | 0.325*** [0.0106] | 0.342*** [0.0109] |
| Experience >30 years | 0.459*** [0.0185] | 0.420*** [0.0180] | 0.437*** [0.0188] | 0.411*** [0.0190] | 0.387*** [0.0120] | 0.403*** [0.0121] | 0.391*** [0.0117] | 0.343*** [0.0117] | 0.339*** [0.0115] | 0.345*** [0.0116] |
| College | 0.324*** [0.0133] | 0.387*** [0.0136] | 0.448*** [0.0149] | 0.462*** [0.0151] | 0.486*** [0.00938] | 0.521*** [0.00964] | 0.522*** [0.00920] | 0.534*** [0.00923] | 0.511*** [0.00922] | 0.555*** [0.00950] |
| Technical school | 0.189*** [0.0166] | 0.236*** [0.0172] | 0.276*** [0.0194] | 0.249*** [0.0196] | 0.255*** [0.0121] | 0.274*** [0.0125] | 0.276*** [0.0120] | 0.270*** [0.0124] | 0.266*** [0.0124] | 0.274*** [0.0130] |
| Senior high school | 0.104*** [0.0134] | 0.142*** [0.0139] | 0.144*** [0.0155] | 0.179*** [0.0154] | 0.162*** [0.00999] | 0.161*** [0.0102] | 0.163*** [0.00986] | 0.171*** [0.0101] | 0.151*** [0.0100] | 0.170*** [0.0103] |
| Provincial dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 9002 | 8861 | 8662 | 8490 | 20926 | 23667 | 25002 | 25878 | 26149 | 24240 |
| R-squared | 0.30 | 0.27 | 0.27 | 0.26 | 0.23 | 0.24 | 0.25 | 0.23 | 0.23 | 0.26 |

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Selected regression results from the extended model, 1988-2007

| | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
|------------------------|-----------------------|-----------------------|----------------------|------------------------|------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| State sector emp. | 0.125*** [0.0336] | 0.018 [0.0343] | 0.043 [0.0341] | 0.162*** [0.0352] | -0.070** [0.0293] | -0.137*** [0.0291] | -0.110*** [0.0259] | -0.081*** [0.0257] | -0.095*** [0.0241] | -0.090*** [0.0216] |
| Experience 11-20 years | 0.360*** [0.0109] | 0.346*** [0.0119] | 0.360*** [0.0109] | 0.284*** [0.0106] | 0.269*** [0.00951] | 0.290*** [0.0114] | 0.272*** [0.0128] | 0.280*** [0.0131] | 0.283*** [0.0136] | 0.267*** [0.0146] |
| Experience 21-30 years | 0.480*** [0.0118] | 0.467*** [0.0124] | 0.493*** [0.0110] | 0.419*** [0.0108] | 0.404*** [0.00950] | 0.440*** [0.0112] | 0.443*** [0.0130] | 0.418*** [0.0132] | 0.409*** [0.0135] | 0.395*** [0.0143] |
| Experience >30 years | 0.546*** [0.0130] | 0.539*** [0.0135] | 0.577*** [0.0121] | 0.495*** [0.0123] | 0.477*** [0.0110] | 0.500*** [0.0131] | 0.511*** [0.0153] | 0.488*** [0.0153] | 0.496*** [0.0158] | 0.458*** [0.0171] |
| College | 0.124*** [0.0115] | 0.146*** [0.0118] | 0.163*** [0.0104] | 0.169*** [0.00998] | 0.197*** [0.00860] | 0.241*** [0.0101] | 0.308*** [0.0116] | 0.256*** [0.0116] | 0.272*** [0.0119] | 0.261*** [0.0126] |
| Technical school | 0.0659*** [0.0130] | 0.0818*** [0.0139] | 0.121*** [0.0121] | 0.122*** [0.0116] | 0.128*** [0.0103] | 0.151*** [0.0121] | 0.194*** [0.0141] | 0.166*** [0.0143] | 0.173*** [0.0146] | 0.157*** [0.0159] |
| Senior high school | 0.0187* [0.0103] | 0.0157 [0.0108] | 0.0141 [0.00955] | 0.0390*** [0.00945] | 0.0473*** [0.00856] | 0.0707*** [0.0101] | 0.0890*** [0.0119] | 0.0687*** [0.0117] | 0.0844*** [0.0119] | 0.0918*** [0.0126] |
| Provincial dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7729 | 6941 | 7317 | 7435 | 9648 | 9288 | 9254 | 9170 | 9220 | 9184 |
| R-squared | 0.347 | 0.33 | 0.387 | 0.377 | 0.383 | 0.373 | 0.361 | 0.335 | 0.334 | 0.317 |
| | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| State sector emp. | -0.005 [0.0206] | 0.006 [0.0184] | -0.007 [0.0174] | 0.045*** [0.0168] | 0.143*** [0.00913] | 0.180*** [0.00895] | 0.204*** [0.00814] | 0.227*** [0.00789] | 0.205*** [0.00768] | 0.198*** [0.00774] |
| Experience 11-20 years | 0.275*** [0.0156] | 0.251*** [0.0162] | 0.279*** [0.0172] | 0.272*** [0.0182] | 0.239*** [0.0109] | 0.270*** [0.0112] | 0.281*** [0.0108] | 0.238*** [0.0105] | 0.264*** [0.0106] | 0.256*** [0.0110] |
| Experience 21-30 years | 0.398*** [0.0152] | 0.361*** [0.0159] | 0.387*** [0.0169] | 0.374*** [0.0178] | 0.306*** [0.0107] | 0.322*** [0.0110] | 0.305*** [0.0105] | 0.271*** [0.0104] | 0.289*** [0.0105] | 0.301*** [0.0109] |
| Experience >30 years | 0.460*** [0.0185] | 0.419*** [0.0182] | 0.438*** [0.0191] | 0.402*** [0.0193] | 0.353*** [0.0121] | 0.357*** [0.0122] | 0.338*** [0.0117] | 0.278*** [0.0117] | 0.284*** [0.0116] | 0.291*** [0.0117] |
| College | 0.324*** [0.0134] | 0.386*** [0.0138] | 0.449*** [0.0151] | 0.455*** [0.0154] | 0.451*** [0.00958] | 0.474*** [0.00983] | 0.469*** [0.00933] | 0.463*** [0.00941] | 0.445*** [0.00943] | 0.487*** [0.00975] |
| Technical school | 0.189*** [0.0166] | 0.235*** [0.0173] | 0.277*** [0.0196] | 0.244*** [0.0197] | 0.233*** [0.0122] | 0.240*** [0.0125] | 0.239*** [0.0120] | 0.226*** [0.0123] | 0.226*** [0.0123] | 0.235*** [0.0129] |
| Senior high school | 0.104*** [0.0134] | 0.142*** [0.0139] | 0.144*** [0.0155] | 0.177*** [0.0154] | 0.149*** [0.00997] | 0.143*** [0.0102] | 0.144*** [0.00977] | 0.150*** [0.00996] | 0.130*** [0.00994] | 0.147*** [0.0102] |
| Provincial dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 9002 | 8861 | 8662 | 8490 | 20926 | 23667 | 25002 | 25878 | 26149 | 24240 |
| R-squared | 0.296 | 0.274 | 0.265 | 0.257 | 0.24 | 0.255 | 0.266 | 0.255 | 0.253 | 0.275 |

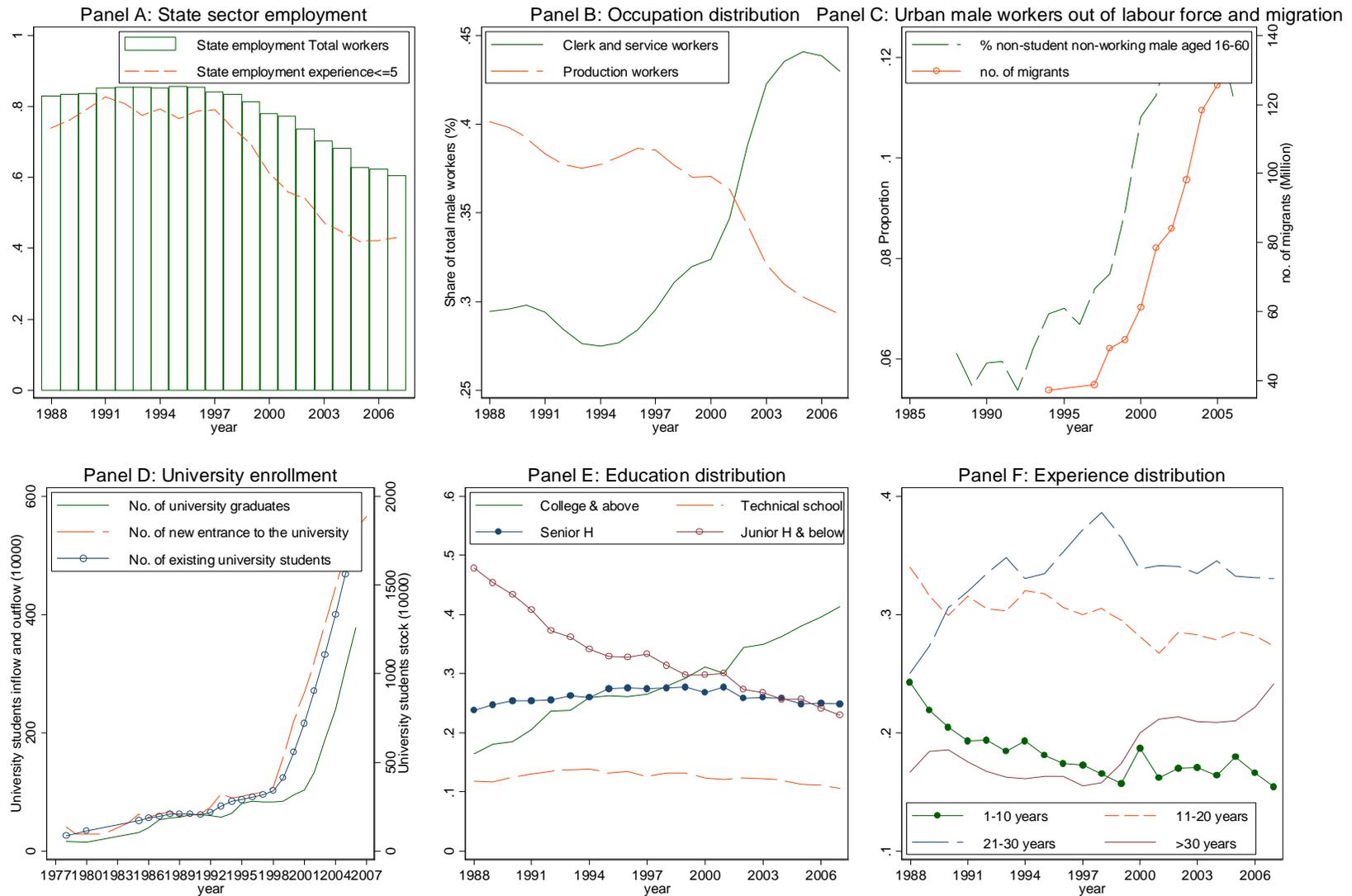
*** p<0.01, ** p<0.05, * p<0.1

Standard errors in brackets

Table 5: Decomposition of the variance of log annual earnings

| | <u>2007-1988</u> | | <u>1993-1988</u> | | <u>1998-1993</u> | | <u>2003-1998</u> | | <u>2007-2003</u> | |
|------------------------------------|------------------|------|------------------|-------|------------------|------|------------------|------|------------------|-------|
| | Mean | % | Mean | % | Mean | % | Mean | % | Mean | % |
| Basic model, 1988 as the base year | | | | | | | | | | |
| Total actual difference | 0.204 | | 0.024 | | 0.109 | | 0.076 | | -0.004 | |
| Observable price difference | 0.046 | 23% | 0.021 | 89% | 0.020 | 18% | 0.004 | 5% | 0.002 | -40% |
| Composition | -0.027 | -13% | -0.019 | -78% | -0.014 | -13% | -0.009 | -12% | 0.015 | -328% |
| Within variance | 0.185 | 91% | 0.021 | 89% | 0.103 | 95% | 0.081 | 107% | -0.021 | 468% |
| Basic model, 2007 as base year | | | | | | | | | | |
| Observable price difference | 0.031 | 15% | 0.012 | 49% | 0.021 | 19% | 0.002 | 3% | -0.003 | 68% |
| Composition | -0.008 | -4% | 0.005 | 21% | -0.008 | -7% | -0.007 | -10% | 0.002 | -55% |
| Within variance | 0.180 | 88% | 0.007 | 30% | 0.096 | 88% | 0.081 | 107% | -0.004 | 87% |
| Extended model, 2007 as base year | | | | | | | | | | |
| Observable price difference | 0.026 | 12% | 0.014 | 57% | 0.017 | 15% | -0.003 | -4% | -0.005 | 111% |
| Composition | 0.047 | 23% | 0.056 | 234% | -0.002 | -2% | 0.038 | 51% | 0.021 | -473% |
| Within variance | 0.136 | 65% | -0.046 | -191% | 0.094 | 87% | 0.040 | 53% | -0.021 | 462% |

Figure 1: Various aspect of urban labour supply and demand



Notes: Data for Panel A, B, E, and F are for 16 provinces using Urban Household Income and Expenditure Survey (UHIES) 1986-2007. For Panel C, the series for the non-labour force is from UHIES and the series for migration is from World Bank (2009). For Panel D the data are obtained from Educational Statistical Yearbook of China (Various years).

Figure 2: Distribution of male earnings

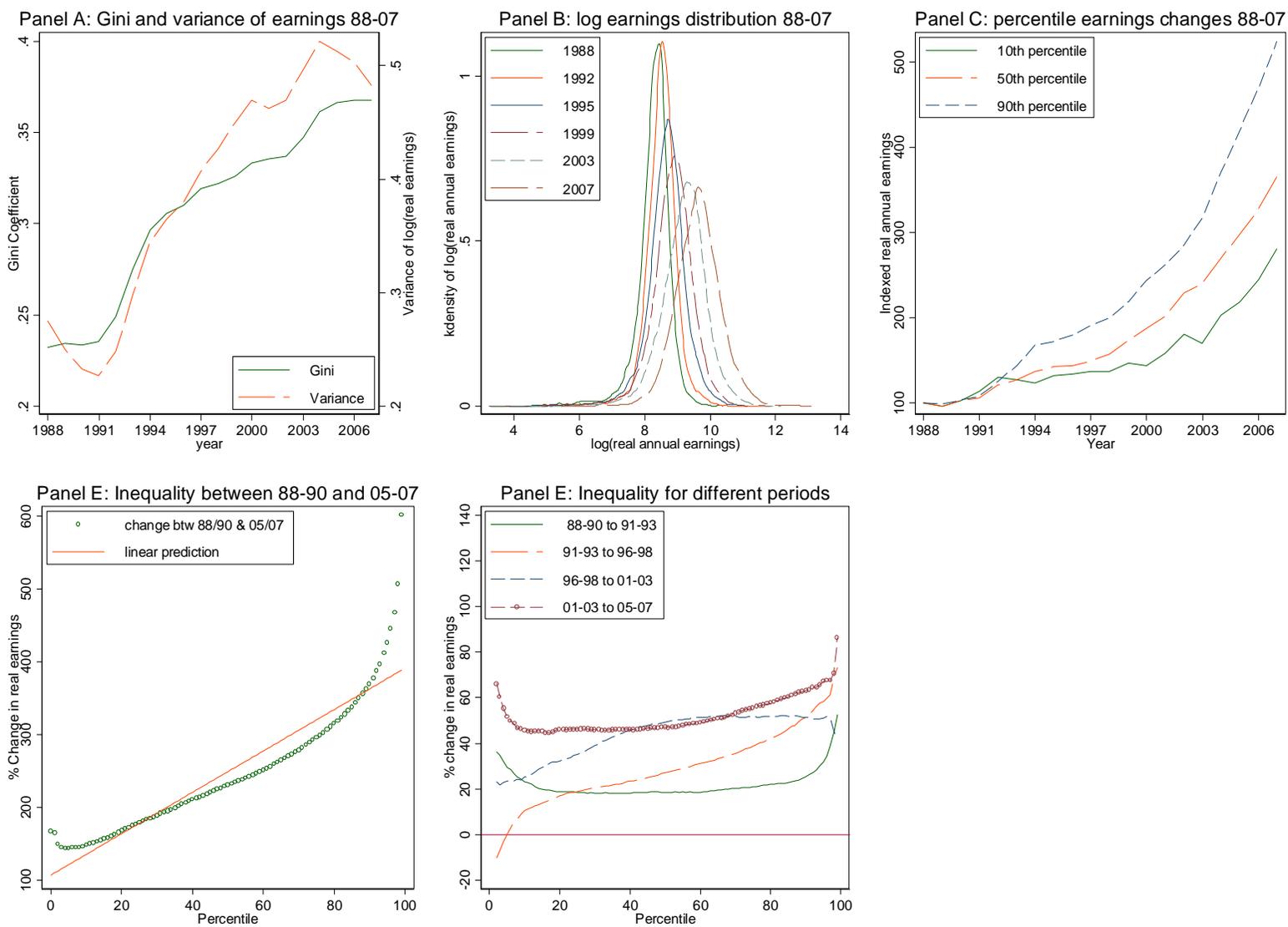


Figure 3: Variance of log earnings for subgroups

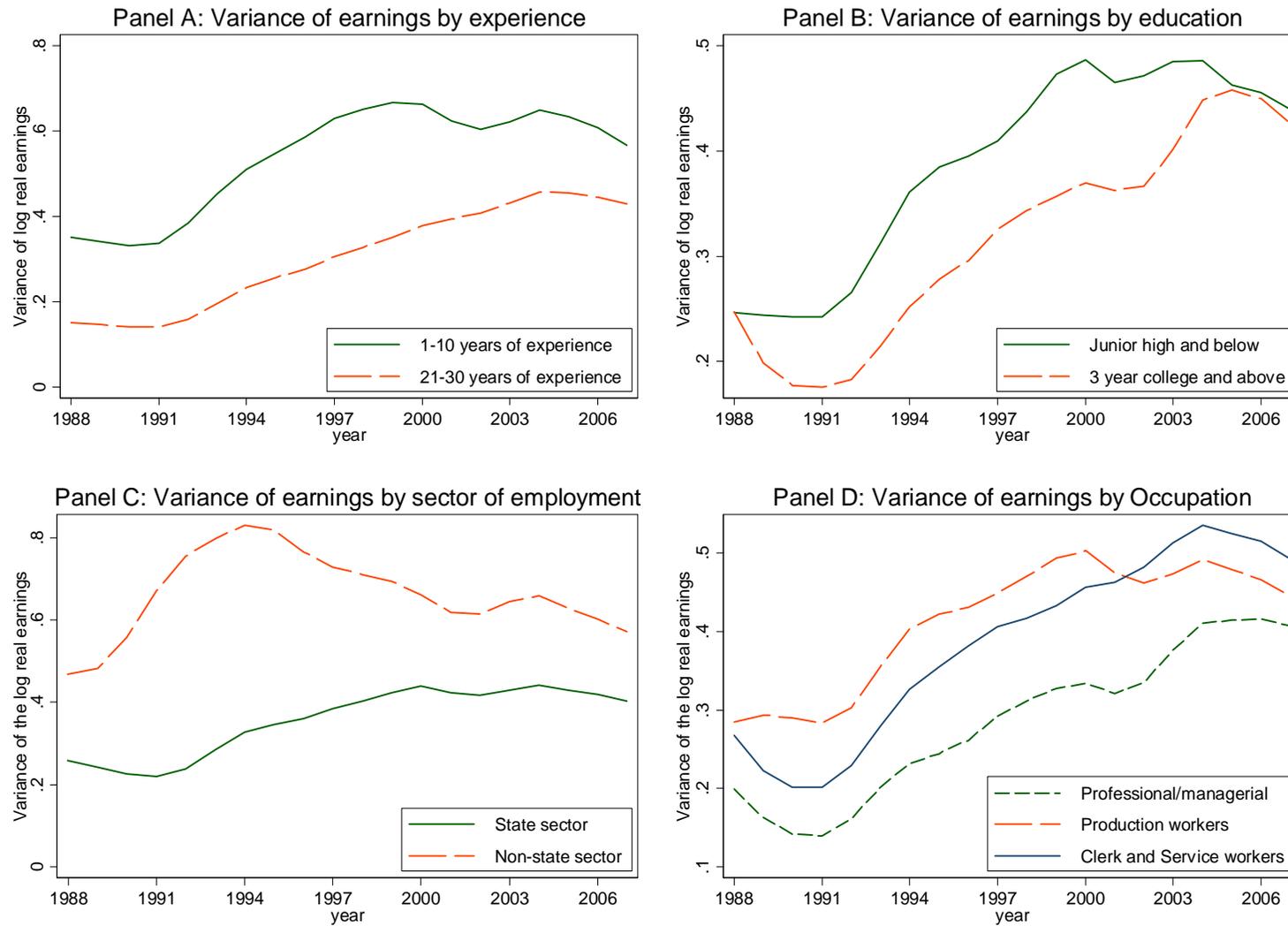


Figure 4: Change in returns to observable characteristics

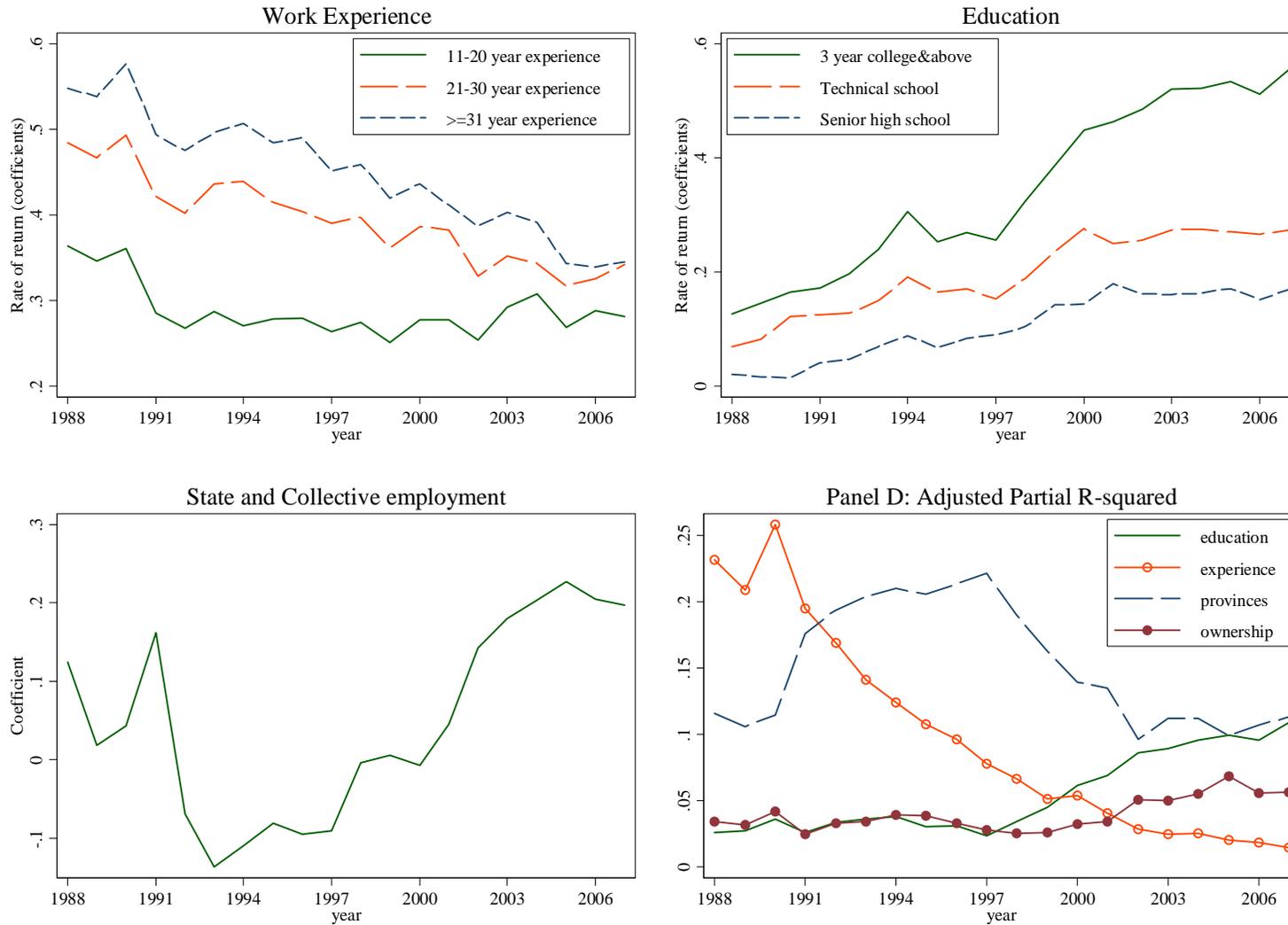
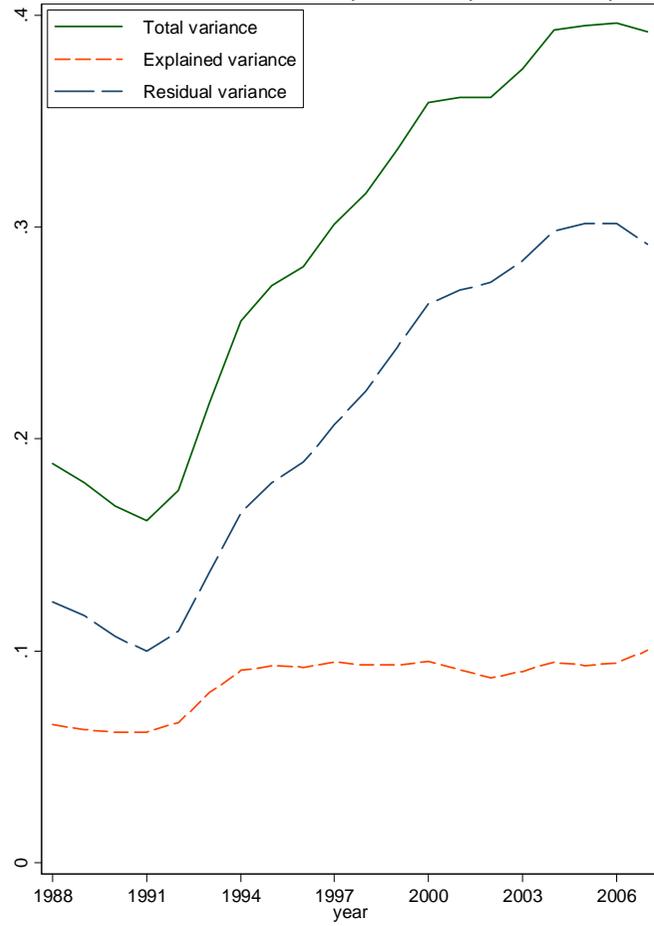


Figure 5: Variance decomposition (Blinder-Oaxca decomposition)

Panel A: Total variance decomposition explained/unexplained



Panel B: Decomposition of explained portion Xs or Betas

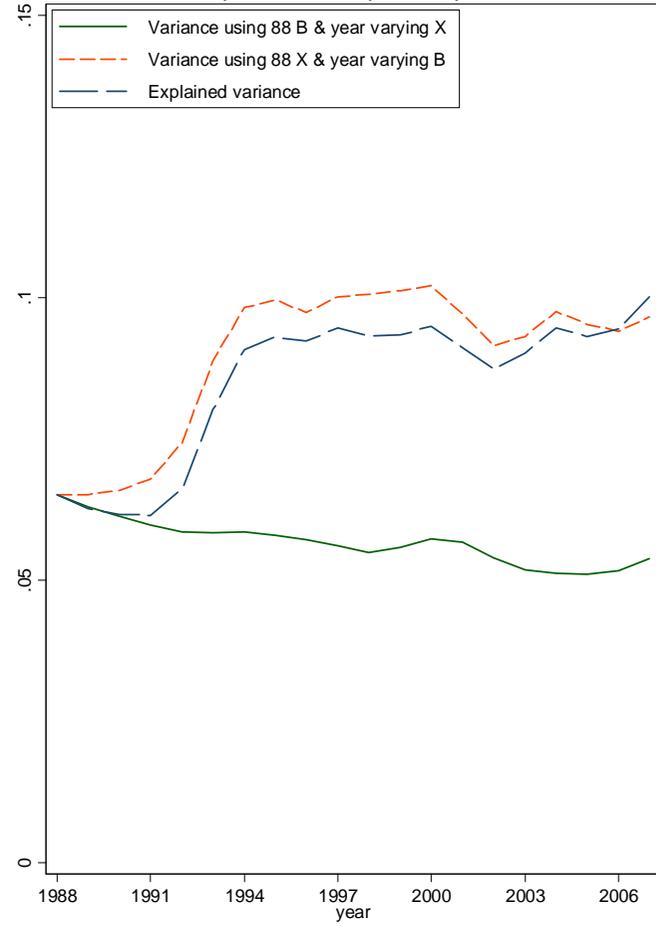


Figure 6: Decomposition of the change in log earnings variance

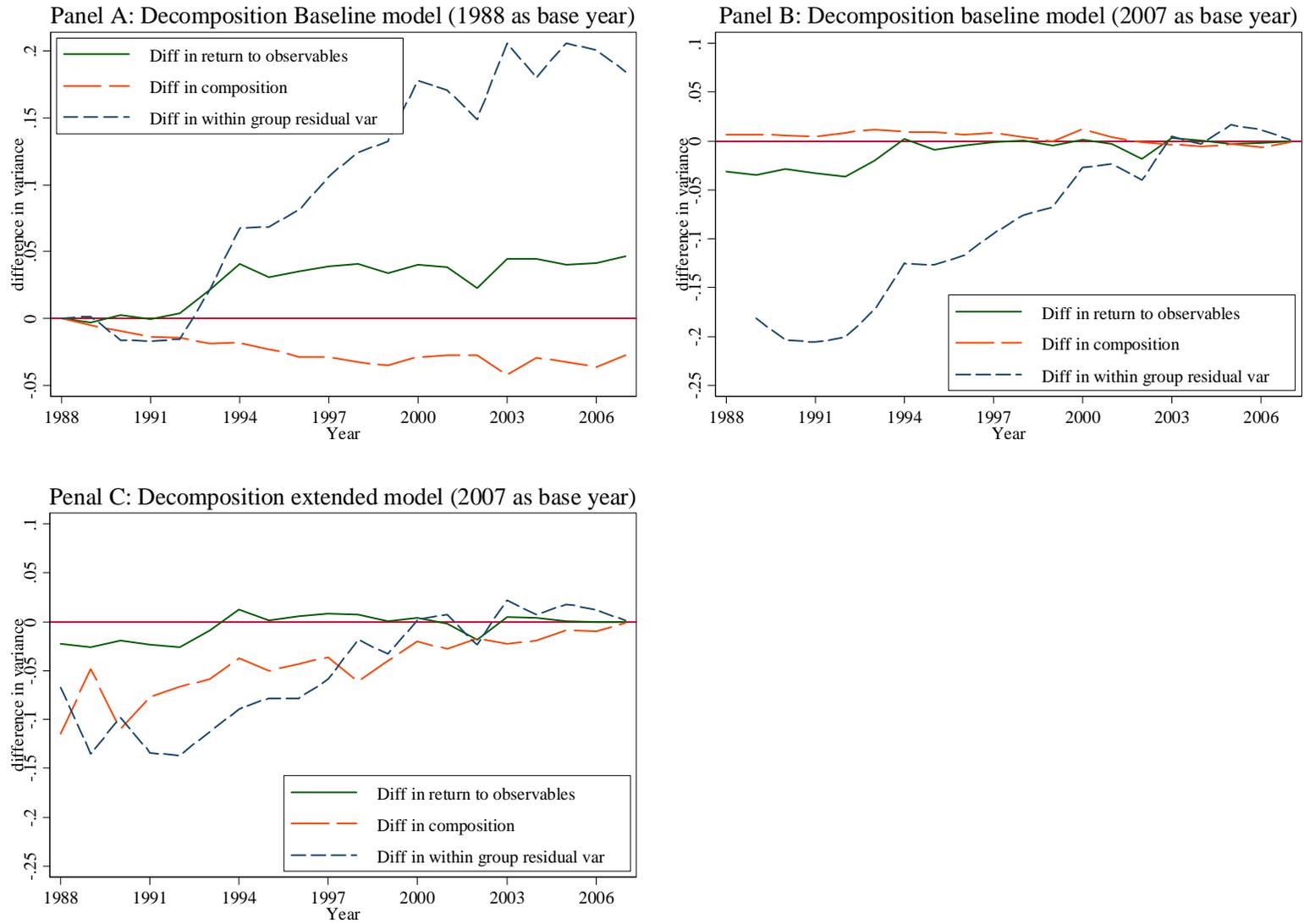


Figure 7: Illustration of the impact of college expansion on change in the distribution of unobserved skills

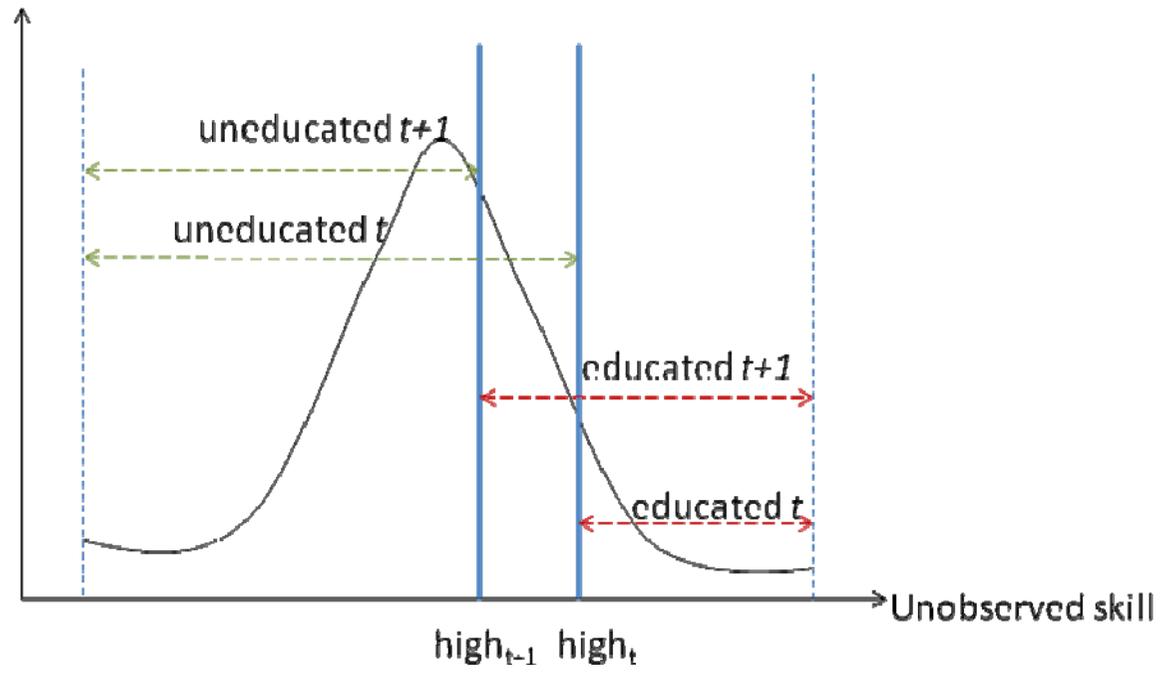


Figure 8: Within group residual variance change

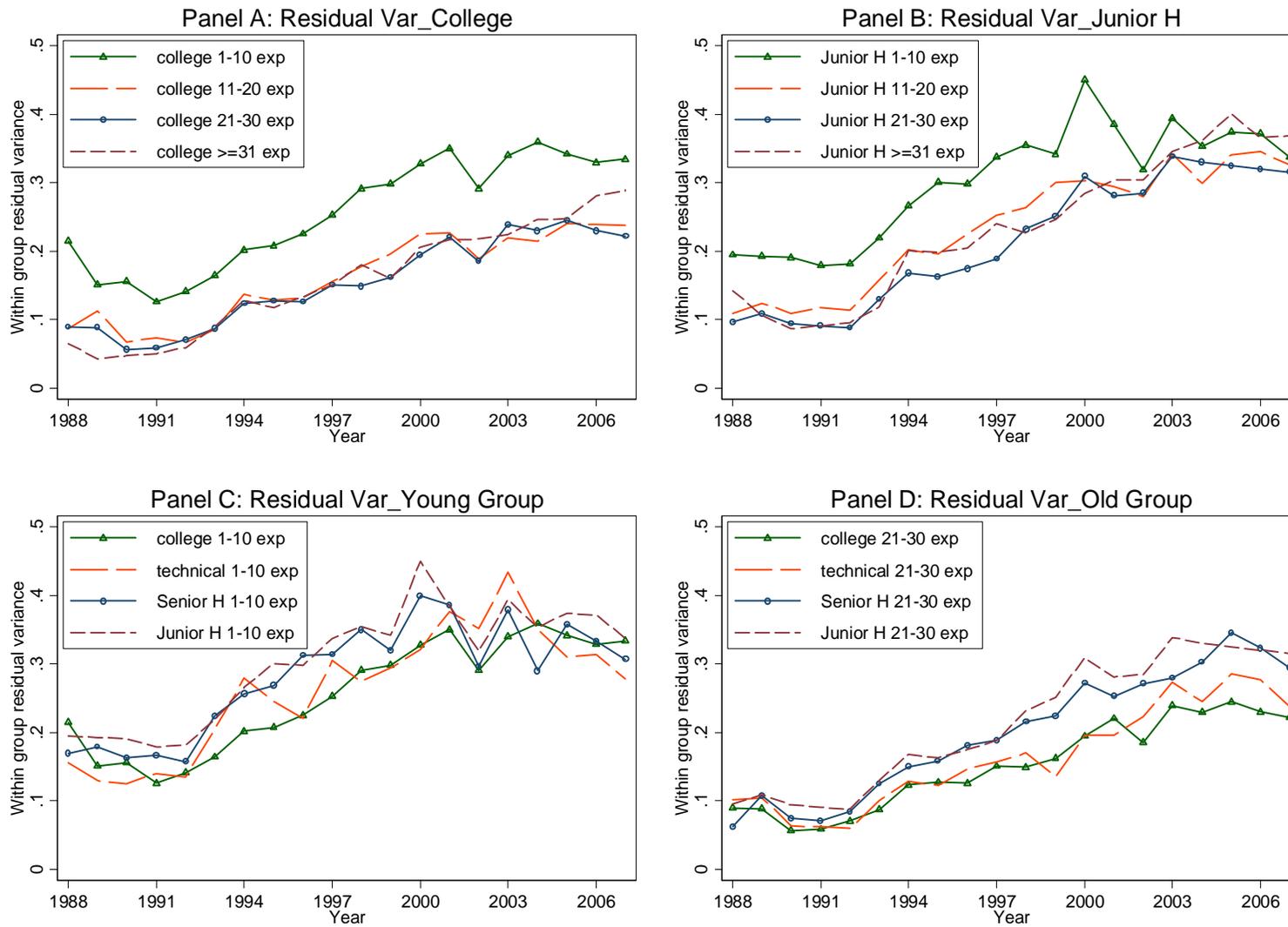


Figure 9: Reproduction of Panels C and D of Figure 8

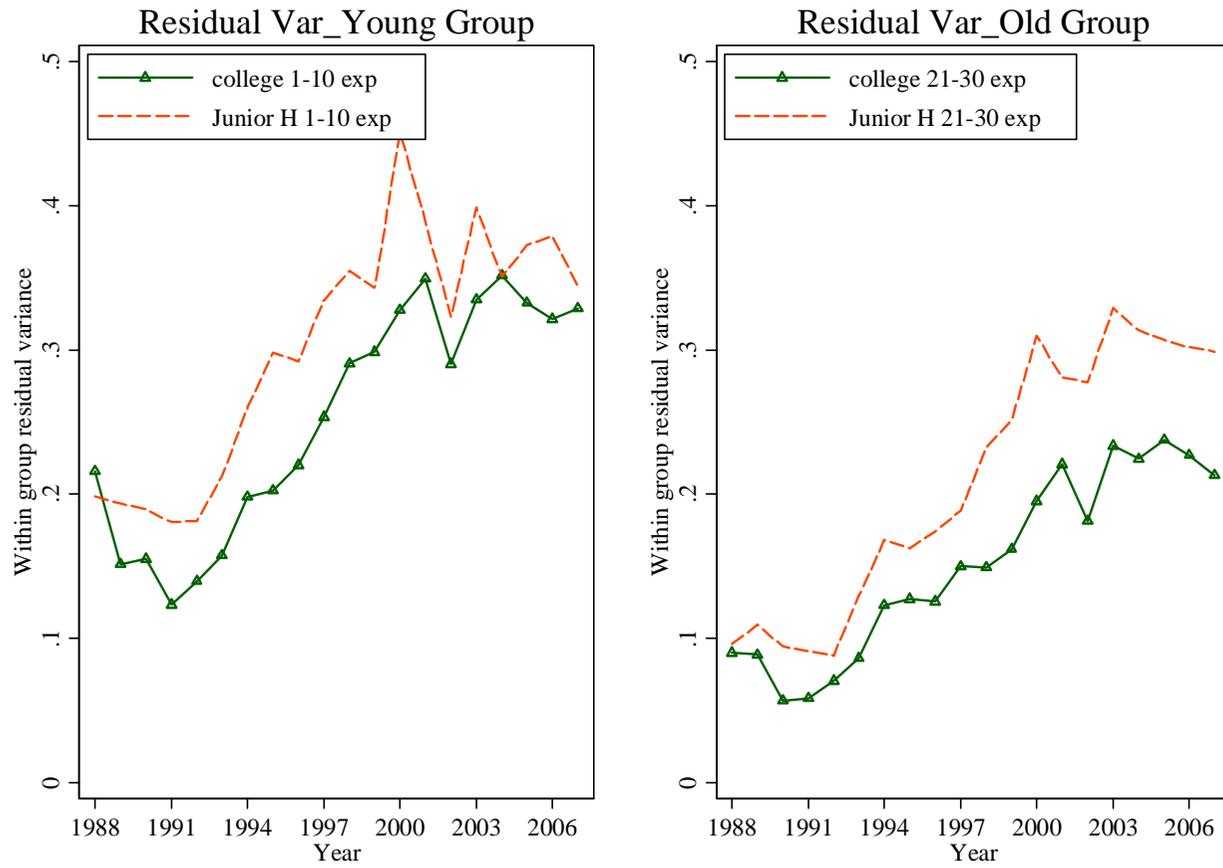
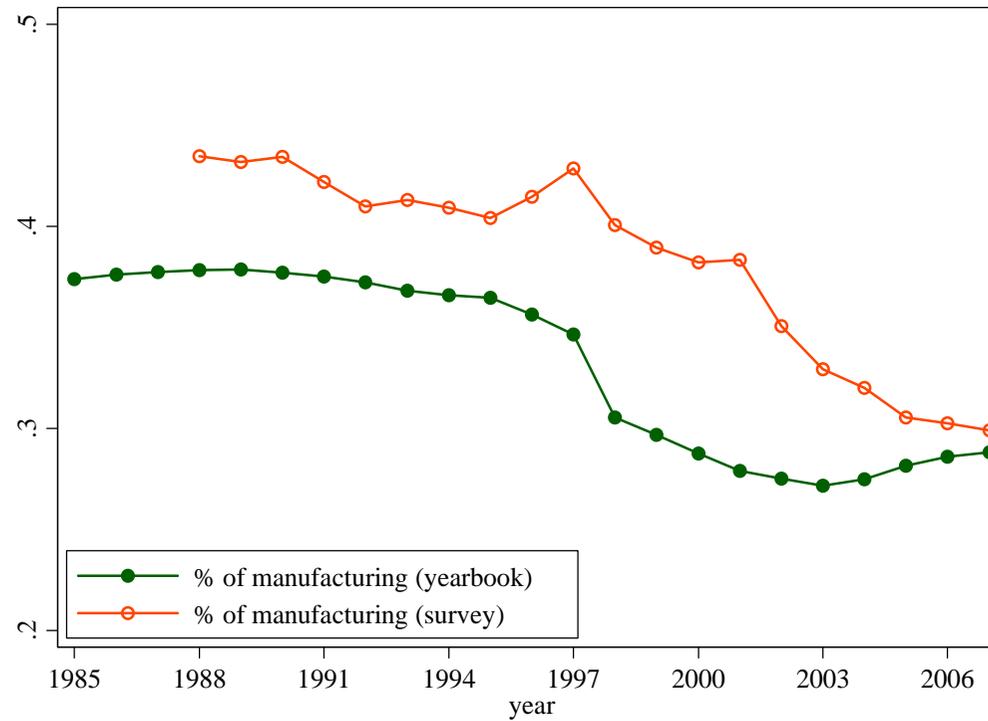


Figure 10: Share of urban workforce in the manufacturing sector



Note: Yearbook data are from China Statistical Yearbooks 2000, 2004 and 2008 (NBS, various years).

Figure 11: Difference-in-differences of within group residual variances

