



Via Po, 53 – 10124 Torino (Italy)  
Tel. (+39) 011 6704043 - Fax (+39) 011 6703895  
URL: <http://www.de.unito.it>

## WORKING PAPER SERIES

### INCOME CONCENTRATION IN CHINA: WHAT ROLE FOR EDUCATION?

Donatella Saccone

Dipartimento di Economia "S. Cognetti de Martiis"

Working paper No. 02/2012



Università di Torino

# Income Concentration in China: What Role for Education?

Donatella Saccone

Department of Economics “S. Cagnetti de Martiis”  
University of Torino  
Via Po 53  
10124  
Torino - Italy  
E-mail: donatella.saccone@unito.it

## Abstract

The paper aims to analyze the relationship between income concentration and inequality in education in China for the period 1989-2006. We first estimate the association between individual attributes and income over time and, then, calculate the contribution of education and its distribution to income concentration.

We find that education has become the bulk of income concentration and that the pace of equalizing reforms should be accelerated, since the improvements in terms of educational equality are at risk to be counterbalanced by the fast growth of its importance in determining income distribution.

**Keywords:** China, Income inequality, Income concentration, Education, Inequality Measurement.

**JEL Codes:** C43, D63, I21, O15.

## 1. Introduction

A series of scientific works have shown that in thirty years of economic reforms the Chinese income distribution has significantly changed. Although their results partially differ, they agree on a number of significant points.

First, it has been demonstrated that income inequality has rapidly increased and reached a considerably high level when compared to its initial starting point. Until the end of '70s, China was indeed characterized by a low level of economic and social inequality, especially in villages and cities. However, from the mid-'80s to the beginning of the new millennium, income inequality indicators almost doubled (Lee, 2000 ; Huang et al., 2003; Galbraith et al., 2004 ; Wu and Perloff, 2005; Benjamin et al., 2005; Wagstaff, 2005; Chen and Ravallion, 2007). Contemporaneously, people's living standards improved, with the percentage of the poor population changing from 53% to 5% from 1980-2005 (Chen and Ravallion, 2007).

Second, income distribution also worsened at a regional level (Xu and Zou, 2000; Fang et al., 2002; Wu and Perloff, 2005; Benjamin et al., 2005; Feng and Yu, 2006; Chen and Ravallion, 2007). The income gap between rural and urban areas widened, as well as the difference in the per-capita income across provinces. However, the contribution of territorial disparities to the national income distribution has often been overestimated. Recent studies show that the greatest contribution to the national inequality must be ascribed to distributive dynamics within, rather than between, areas and provinces (Lee, 2000; Benjamin et al., 2002; Sicular et al., 2007).

In order to explain the evolution of the Chinese income distribution, a series of causes have been identified. At a national level, it seems that important roles have been played by the rapid, but not equally distributed, economic growth; the raising of new

benefits to the monopolistic sectors such as public, financial, and transport services; the increasing corruption; the decrease in the relative role of SOEs, and the growing inflation rates (Xu and Zou, 2000; Chang, 2002; Galbraith et al., 2004). At a provincial level, biased economic policies, a better availability of infrastructures, and the geographical proximity to foreign markets, such as Hong Kong, Taiwan, Japan, and Korea, have fostered coastal provinces to the detriment of internal areas (Chang, 2002 ; Yao et al., 2004). At the same time, urban-biased policies, disparities in social services, the slowdown in the relative growth of agricultural incomes, and the uneven access to non-agricultural activities have led to the relative backwardness of rural areas (Zhang and Kanbur, 2005; Benjamin et al., 2005; Bardhan, 2010).

A particular attention has been paid to the role of education in influencing income distribution. The skill-biased technological change and the subsequent growth of the demand for more educated workers have worsened wage inequality by increasing the skill-premium, especially in the urban sector (Wu and Xie, 2002; Heckman and Li, 2003; Benjamin et al., 2005; Heckman, 2005; Go et al., 2009; Ning, 2010; Bardhan, 2010). Moreover, in rural areas education represents an important individual asset to develop managerial skills, cope with new economic opportunities, and generate non-agricultural incomes (Benjamin et al., 2002). Sicular et al. (2007) found that education has been the most relevant individual asset explaining the rural-urban gap. They estimated that if the education level in rural areas had been equal to the urban one, the income gap would have decreased by 25-30%. In this regard, Fang et al. (2002, p. 442) stated that “investment in human capital is key to long-term improvements in welfare for all”, while Heckman (2005, p.66) affirmed that “human capital is the asset that ultimately determines the wealth of China. Fostering access to education will reduce inequality in the long run”.

However, no effort has been made to directly relate income distribution and education distribution. Actually, the contribution of education to income inequality is strictly related not only to the association between individual education level and individual income, but also to the way in which education is distributed among individuals. Given the potential role of education in affecting income inequality in China, this paper aims to analyze how income concentration is related to inequality in education by isolating the role of education from the role of the other determinants of income distribution. By using household survey data for the period 1989-2006 and following the economic literature on income-related health inequality, we first estimate the association between individual attributes – such as age, residence, gender, income composition, and education- and income over time, and then calculate the contribution of education and its distribution to income concentration. Paragraph 2 presents data and methodology. In paragraph 3 we report the results of the analysis. In paragraph 4 we discuss the results and provide policy considerations. In the last paragraph, conclusions are drawn.

## **2. Data and methodology**

The main data source on which the analysis is based is the *China Health and Nutrition Survey* (CHNS). Since it provides complete information at both individual and household level across time, it has been largely used in the literature on inequality in China (see for example Zhang and Kanbur, 2005; Li and Zhu, 2006; Goh et al., 2009; Ning, 2010). This is a multistage, random cluster sample survey designed by a group of social scientists and biomedical researchers under the control of the Carolina Population Center, the National Institute of Nutrition and Food Safety, and the Chinese Center for Disease Control and Prevention. The survey was collected in various years (1989-1991-1993-1997-2000-2004-2006), and the Carolina Population Center provides a longitudinal

database favoring inter-temporal and inter-spatial comparability <sup>1</sup>. Even if its focus is on health issues, for each interviewed person the survey also provides the number of completed years of education and the total income at constant price, along with a series of individual attributes (age, province of residence, area of residence). We focus on individuals aged 15 and above, who were not enrolled at school when the survey was collected.

The CHNS covers 9 Chinese provinces: Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, Guizhou. Since the provinces of Liaoning and Heilongjiang were not covered by the survey in 1989 and 1991, we limit our analysis to the remaining 7 provinces. The 7 provinces differ in regards to geography, economic conditions, social indicators, and public resources. Since the richest and poorest provinces are not considered by the survey, it could underestimate the level of heterogeneity characterizing the country. However, each province has a population that is larger than that of many developing countries. For this reason, although conclusions cannot be extended to the whole country, they are significant for a notable number of individuals (Benjamin et al., 2002). Moreover, despite its incomplete geographical coverage, the CHNS is particularly useful for our purposes. On one hand, it allows us to jointly analyze both education and income distribution. On the other hand, given its time coverage, the survey allows us to study the evolution of the two distributions over a period of 17 years.

In order to estimate the relationship between income distribution and inequality in education, we adopt and adjust to our case the methodology frequently used by economic literature on health inequality (Jones and Nicolàs, 2004; Chen et al., 2005;

---

<sup>1</sup> The survey is collected over a 3-day period on a sample of about 4400 households (19000 individuals). However, the longitudinal database is smaller: excluding the provinces of Liaoning and Heilongjiang, we have observations for about 3300 households (9300 individuals). Since we focus on individuals aged 15 and above not enrolled at school when the survey was collected, we have 8403, 8465, 7568, 7588, 7952, 6821, 6619 observations in 1989, 1991, 1993, 1997, 2000, 2004 and 2006 respectively. For details see the website of the Carolina Population Center: [www.cpc.unc.edu/projects/china/data](http://www.cpc.unc.edu/projects/china/data).

Wagstaff et al., 1991 and 2001; Lahiri and Zulkarnain, 2007; O'Donnell et al., 2008). In these studies, a health concentration curve is drawn by plotting the cumulative percentage of the health variable against the cumulative percentage of population ranked by income. Then, the health concentration index is defined as twice the area between the health concentration curve and the 45-degree line (O'Donnell et al., 2008). In turn, “the health concentration index can be decomposed into the contributions of individual factors to income-related health inequality, in which each contribution is the product of the sensitivity of health with respect to that factor and the degree of income-related inequality in that factor” (O'Donnell et al., 2008, p. 159).

We adapt this methodology to our research question –i.e. how much income distribution is related to inequality in education - in three steps. First, we substitute the two variables, health and income, with income and education respectively. In other words, the individual income is used as the main variable, and its cumulative percentage is plotted against the cumulative percentage of population ranked from the least to the most educated individual. The resulting concentration curve represents how income is distributed in relation to the individual level of education. When the concentration curve lies on the 45-degree line, income is equally distributed across the education levels, and the concentration index is equal to zero. The more the income concentration curve lies below the 45-degree line, the more income is concentrated in the hands of the most educated individuals, and the higher is the value of the concentration index.<sup>2</sup>

The concentration index can be defined by the following formula (Wagstaff et al., 2001):

$$C = \frac{2}{n\mu} \sum_{i=1}^n y_i R_i - 1 \quad [1]$$

---

<sup>2</sup> Even if it is an improbable case, the income concentration curve could lie above the diagonal. In this case, the concentration index would be negative and it would mean that income is concentrated in the hands of the least educated individuals.

In our case,  $y_i$  represents the individual income,  $\mu$  is the mean of  $y$  and  $R_i$  is the fractional rank of the  $i^{th}$  person in the education distribution. Sometimes, the concentration index has been criticized because of the difficult interpretation of its values; in particular, since it is not expressed in natural units, it is not simple to ascribe its values to a low or a high degree of inequality (Milanovic, 1997; Koolman and Doorslaer, 2004). However, we are interested in tracing its evolution over time rather than interpreting the meaning of a single value. Moreover, additional information on income concentration will be provided by the decomposition of the index, as illustrated below (see equation 3). Most importantly, the widely recognized advantage of the concentration index is that it is a bivariate measure of inequality; this allows us to measure inequality in income related to the ranking of the individual education level. Indeed, it takes into account not only the coefficient of variation of income, but also the correlation between income and education rank (Milanovic, 1997; Koolman and Doorslaer, 2004).

The concentration curve and the value of the concentration index are necessary, but not sufficient, tools to understand how income distribution is related to inequality in education. Indeed, the income distribution across education levels depends on a number of factors, in particular on the sensitivity of income to other individual attributes, and the concentration of those attributes across education levels. In other words, education concurs to influence individual income along with other individual attributes. The concentration curve reflects the income inequality directly related to education, as well as the income inequality related to other attributes differing across education levels and influencing individual income. For example, people with a low educational level could live in geographical areas in which on average per-capita income as well as education are lower. In this case, even if education did not have any impact on the income level, this would result in a high value of education-related income concentration. However, this value would reflect the fact that less educated individuals live in areas in which the

income levels are lower, regardless of the level of education, rather than the direct effect of education on individual income. At the same time, in the extreme case of the income being equally distributed (all the individuals have the same income), income would also be equally distributed also across education levels. In other words, income concentration is the result of inequality in education as well as other determinants. From this point of view, income concentration is the result of income inequality across education levels and income distribution across the other determinants of individual income.

In order to estimate the contribution of these determinants to income concentration and to isolate the effect of education, the second step of our analysis is to define the individual income as a linear function of a vector of individual attributes  $X_i$  by adapting the methodology proposed by Walgstaff et al. (1991)<sup>3</sup>:

$$y_i = \alpha + \beta X_i + \varepsilon_i \quad [2]$$

The model is estimated by using OLS and the vector of individual characteristics includes<sup>4</sup>:

- age and age squared to typify possible non-linear effects
- two dummy variables, urban and male, to consider income differences by gender and area
- education, measured by the number of completed years of schooling
- the percentage of each income source (agriculture, business and wage) on the total individual income to reflect inequality across sources of income
- dummies for province of residence to take into account geographical inequality.

---

<sup>3</sup> The literature on health inequality expresses the health variable as a function of a vector of individual attributes.

<sup>4</sup> It must be underlined that the estimated coefficients represent a measure of association between income and the explanatory variables, rather than a relationship of causality. As a consequence, we just estimate how income is associated or related to some individual factors.

In order to link individual income and its determinants to income concentration, the third step consists of rewriting the concentration index as the weighted sum of the concentration index of each independent variable (O'Donnell et al., 2008), where weights are represented by the elasticity of income with respect to that variable:

$$C = \sum_k (\beta_k \bar{x}_k / \mu) C_k + GC_\varepsilon / \mu \quad [3]$$

where  $\mu$  is the average income,  $\bar{x}_k$  is the mean of each regressor and  $C_k$  is the concentration index of that regressor (i.e. the concentration of the regressor across educational levels). The second term, calculated as a residual, represents the part of the concentration index that is not explained by our specification. The lower is the residual, the higher is the explanatory power of the specification. In this way, the contribution of each individual attribute to the income concentration index depends on two components. The first is the elasticity of income with respect to that attribute. The second is the concentration of that attribute across education levels or, in other words, the concentration index of that explanatory variable. Notice that when the explanatory variable is the individual education level, its concentration index is calculated by plotting the cumulative percentage of education against the cumulative percentage of the population ranked by education level, which is by definition the education Gini coefficient corresponding to the generalized Lorenz curve.

### 3. Education-related income inequality in China

Figure 1 plots the income concentration curves from 1989 to 2006. Over this period, the distance between the curve and the equality line (45-degree line) widened.

This reveals that a growing share of income became more and more concentrated in the hands of the most educated individuals, and it is confirmed by the values of the income concentration index, that moved from a low value of 0.09 in 1989 to a relatively high value of 0.24 in 2006 (table 1).

[figure 1 about here]

In order to understand which have been the main determinants of the individual income, how much the latter has been related to the individual education level, and how this relationship has changed over time, we express the individual income as a linear function of a vector of individual attributes: age, area of residence (urban or rural), gender, province of residence, relative composition of income by sources (agriculture, business and wage), and education level. Table 1 shows the results of the OLS regressions, run separately for three different years: 1989 (column 2), 1997 (column 3) and 2006 (column 4).

[table 1 about here]

The nonlinear effect of age is confirmed: in all three years and, especially in 1997 and 2006, individual income increased with age but at a decreasing rate. The reversing effect of age is negligible, and the turning point is estimated to be around 65 years. This could be partially explained by the concentration of high-rank employment positions in the hands of the old elite of workers (Gustafsson and Sai, 2009). The effect of the area of residence has changed over time. In 1989, the coefficient of the urban dummy variable was negative but not significant. On the contrary, in 1997, and especially in 2006, individuals living in urban areas had on average higher income. This reversing tendency

probably reflects the urban-bias of reforms in the '90s (Bardhan, 2010). The gender seems to have a large and highly significant relation with individual income. Not only did males have higher income on average in all the three periods, but this effect became larger over time. The growing importance of territorial differences is also confirmed by the coefficients of the provincial dummy variables. Although not always significant, the signs of the coefficients differed from province to province, and their value became larger over time. The percentage composition of the individual income played a minor role in respect to the other individual attributes. In any case, results suggest that in 1997 and 2006 the total individual income on average was lower when largely based on agricultural activities.

As expected, the coefficient of education is positive, although significant only in 1997 and 2006. It seems that in 1989 the individual education level had a negligible relationship with the individual income. On the contrary, since the '90s it turns out to be a fundamental factor, probably reflecting the growing labor demand for skilled people and the subsequent increase in their rewards (Wu and Xie, 2002; Heckman and Li, 2003; Li, 2003; Benjamin et al., 2005; Heckman, 2005; Go et al., 2009; Ning, 2010; Chi et al., 2011; Chi et al., 2011; Xing and Li, 2011). Moreover, Gustaffson and Sai (2009) show that in 1995 and 2002 education, along with age and gender, was strictly related to the probability of being employed as a worker of high rank.

In order to link the determinants of individual income to the income distribution, we decompose the income concentration index as illustrated in paragraph 2. The results of the decomposition and the relative contribution of each determinant are reported in table 2.

[table 2 about here]

They not only confirm the results of previous literature, i.e. the worsening of the Chinese income distribution (see the values of the income concentration index), but also provide further insights into unveiling the disequalizing factors, the knowledge of which is necessary to design effective equalizing policies. It is clear that in 1989 factors affecting income concentration were notably different from those in 1997 and 2006. They can be synthesized as follows. First, inequality in education is the main determinant of income concentration. Its contribution, already high at the beginning of the period, turned out to be the bulk of income inequality in 2006, accounting for 74% of income concentration. This was the result of two opposite forces. On one hand, the concentration index of education declined over time (from 0.42 in 1989 to 0.32 in 2006) and this, *ceteris paribus*, had an equalizing effect on income concentration. Indeed, if the elasticity of income to education had been equal to the 2006 value across the three years, income concentration would have been higher in 1989 and 1997 than in 2006. It is simple to calculate that in this case the income concentration index would have been equal to 0.30 and 0.28 in 1989 and 1997, respectively, compared to 0.24 in 2006. On the other hand, however, the elasticity of income to education increased, meaning that differences in income became more and more related to education distribution, and that educational inequality become more and more bad for an even distribution of income. As a result, the effect of a more equitable distribution of education was counterbalanced by its growing role in affecting individual income and, then, income distribution.

Second, the percentage contribution of the area of residence (urban or rural) to income concentration increased over time. People living in urban areas, indeed, are concentrated in the highest part of education distribution and on average have higher income (except in 1989). This turned out to be the second determinant of income concentration in 2006, accounting for 12.5%.

Third, the correlation between individual income and gender strengthened between 1989 and 2006, while the distribution of gender across education levels became a little more equitable. As a result, in 2006 the percentage contribution of gender differences to income concentration was still remarkable, representing 6% of income concentration.

Fourth, in 1989 inequality across income sources was the most important factor related to income inequality. In particular, this mostly depended on the fact that individuals whose income was mainly based on wages were the most educated (concentration index of 0.32) and on average had higher income. However, both the effects, although still determinant and true, decreased their importance over time. This explanation probably relies on the structural changes of the Chinese economy over the last 20 years, and especially in the new millennium (Valli and Saccone, 2009). Even if the skill premium grew with the modernization and the adoption of new technologies, it is possible that the average schooling of industrial workers decreased with the massive movement of labor from primary to secondary and tertiary sectors. Moreover, the average level of education increased over the last 20 years, moving from 5.6 years of education in 1990 to 7.6 in 2005 (Barro and Lee, 2010). As a consequence, whoever was relatively more educated in 1989 could have been relatively less educated in 2006 and then belong to the lower-middle part of education distribution. In any case, it seems that income composition became less and less important in affecting income concentration, while the individual education level turned to be the most important feature regardless of the source of income. This suggests that education has become a resource *per se* in all the types of activities, determining to which extent an individual has been able to cope with the new economic possibilities.

Finally, inequality across provinces and age cohorts seems to have a minor role in determining income concentration. In particular, after a first period of growing

importance, geographical differences became relatively less relevant in affecting individual income and income concentration. If on the one hand it confirms the results of previous studies showing that geographical inequalities have actually been overestimated in regard to their contribution to total inequality (Benjamin et al., 2002; Benjamin et al., 2005; Sicular et al., 2007), on the other hand it is worth recalling that our results are based on a sample of 7 provinces and, as a consequence, they could underestimate territorial disparities in income and education.

#### **4. Discussion and policy considerations**

The huge increase in income inequality represents one of the most critical problems for the Chinese economy. Social discontent is giving rise to conflicts and unrest, especially in urban areas (Bardhan, 2010), while the Chinese government is becoming aware of the necessity to combine pro-growth policies with equalizing measures. From our analysis it emerges that in recent years income concentration has been strictly related to three individual characteristics: education, gender and area of residence. First, male individuals living in urban areas on average have higher incomes and a better education. Second, and most important, education has become the bulk of income concentration and, as a consequence, it should also be the bulk of new redistributive and equalizing policies. Moreover, the decomposition results have suggested that the pace of reforms should be accelerated since the improvements in terms of educational equality risk to be counterbalanced by the fast growth of its importance in determining income distribution.

In order to design effective policies, the Chinese education distribution and its determinants should be deeply studied in future research; until now, indeed, few studies have only focalized on spatial and geographical distribution of education (see for example

Zhang and Kanbur, 2005; Hannum and Wang, 2006). At a first glance, it seems that, as the income distribution, the distribution of education has rapidly changed since the beginning of economic reforms. In figure 2, we report the Gini coefficient of education calculated on our sample (individuals aged 15 and above). It constantly decreased until 2004, showing a reversing tendency in 2004 and 2006. To explain this tendency we have to look at two different matters. On one hand, at a national level, the percentage of illiterate population significantly declined from 35.5% in 1975 to 6.5% in 2010, while the average years of schooling moved from 4 to 8 (Barro and Lee, 2010). This resulted in a decline in both educational poverty and educational inequality, at least until 2004. To understand how much the decrease of the Gini coefficient of education has depended on the decrease of the share of illiterate people, on the other hand, we calculate it for a subsample of individuals with at least 6 years of schooling (figure 2). It represents the educational inequality among individuals with at least a basic level of education and allows us to typify what happened to the middle and highest part of the education distribution. What emerges is a slight tendency of the education Gini coefficient to increase. In other words, it seems that the most notable distributional progress has regarded the lowest part of the education distribution, through a more equitable access to basic education, while inequalities in the access to the highest levels of education has held over. This latter effect could partially explain the upward trend of the education Gini coefficient for the whole sample in 2004 and 2006: a more equitable access to basic education has been counterbalanced by an uneven access to the highest school levels. These findings call for additional research as well as new policies. Redistributive policies should be addressed to further spread basic education, and at the same time, to guarantee a more equitable access to the highest levels of education, especially for women and people living in rural areas. The decomposition of income concentration, indeed, has revealed that the highest part of education distribution is overrepresented by male

individuals living in urban areas (see concentration indexes of table 2 for the variables ‘male’ and ‘urban’).

[figure 2 about here]

Until now, urban-biased policies have played an important role in giving rise to urban-rural disparities in education, especially in the '80s and '90s. In the '80s the fiscal reforms decentralizing the funding of education made local governments responsible for its provision. As a consequence, in the poorest areas the public provision of education decreased and its costs were privatized. While private schools were not allowed until the '80s, in the '90s private educational institutions spread, and a process of marketization of education occurred, especially for higher education (Mok, 2000). Moreover, the ‘commune schools’ promoting literacy in rural areas were closed and replaced by schools overseen by higher government-levels, while social services were concentrated in urban areas (Zhang and Kanbur, 2003; Hannum and Wang, 2006). This has been aggravated by the household registration (*hukou*) system, by which households’ eligibility for public urban services is determined (Liu, 2005; Chan, 2010). According to this system, free access to urban primary and secondary schools is available only to families with urban registration, while government rarely grants urban registration to families migrating from rural to urban areas (Yang, 1999; Wu and Treiman, 2004).

More has been done to build new policies reducing education inequality by gender. The “National Guidelines for the Development of Women (2000-2010)” introduced aims and policies to promote gender equality. In particular, a series of specific objectives have been listed regarding the education of women, among which is the full right for girls to receive nine years of compulsory education and the upgrade of girls’ gross enrollment rate at upper secondary level to 75%, and 15% for tertiary colleges

(Zhang and Gao, 2003). However, the time needed to fulfill these objectives is still long and new policies should be designed with the support of new and deep research on education distribution.

## 5. Conclusions

The aim of the paper was to analyze how income concentration in China is related to inequality in education. To this purpose, our analysis followed three steps. First, we calculated the income concentration index across education levels for 1989, 1997, and 2006. Second, for each of the three years, we expressed the individual income as a linear function of a vector of individual attributes: age, area of residence (urban or rural), gender, province of residence, relative composition of income by sources (agriculture, business and wage), and education level. Third, we calculated the contribution of each individual attributes to income concentration. The main findings of our analysis can be synthesized as follows. First, in recent years income concentration has been strictly related to inequality across education levels, gender, and areas (urban/rural). Second, and most importantly, education has become the bulk of income concentration, accounting for 74% of it, and a growing share of income has been more and more concentrated in the hands of the most educated individuals. As a consequence, education should also become the bulk of new research supporting the design of redistributive and equalizing policies, that should be especially addressed to females and people living in urban areas. Finally, the decomposition results suggested that the pace of reforms should be accelerated since the improvements in terms of educational equality are at risk to be counterbalanced by the fast growth of its importance in determining income distribution. Indeed, in recent years the effect of a more equitable distribution of education has been defeated by its growing role in affecting individual income and, then, income distribution.

## References

Bardhan, P. (2010). *Awakening Giants, Feet of Clay: Assessing the Economic Rise of China and India*. Princeton, New Jersey: Princeton University Press.

Barro, R.J., Lee, J. (2010). *A New Data Set of Educational Attainment in the World, 1950-2010*. NBER Working Paper, 15902.

Benjamin, D., Brandt, L., Giles, J., Wang, S. (2005). *Income Inequality During China's Economic Transition*. Working Papers Tecipa, 238, University of Toronto, Department of Economics.

Benjamin, D., L. Brandt, P. Glewwe, Guo, L. (2002). *Markets, Human Capital, and Inequality: Evidence from Rural China*, in Freeman, R., *Inequality Around the World*. London: Palgrave, pp. 87-127.

Chan, K.W. (2010). The Household Registration System and Migrant Labor in China: Notes on a Debate. *Population and Development Review*, 36(2), 357–364.

Chang, G.H. (2002). The Cause and Cure of China's Widening Income Disparity. *China Economic Review*, 13, 335-340.

Chen, S., Ravallion, M. (2007). China's (Uneven) Progress Against Poverty. *Journal of Development Economics*, 82, 1-42.

Chen, Z., Eastwood, D.B., Yen, S.T. (2005). *Childhood Malnutrition In China: Change Of Inequality In A Decade*. Annual meeting, July 24-27, Providence, RI 19205, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association).

- Chi, W., Li, B., Yu, Q. (2011), Decomposition of the Increase in Earnings Inequality in Urban China: A Distributional Approach. *China Economic Review*, 22, 299-312.
- Fang, C., Zhang, X., Fan, S. (2002). Emergence of Urban Poverty and Inequality in China: evidence from Household Survey. *China Economic Review*, 13, 430-443.
- Feng, J., Yu, Y. (2006). *Income Inequality and Health in Rural China*. Paper presented at 2006 UNU/WIDER Conference “Advancing Health Equity”, 29-30 September.
- Galbraith, J.K., Krytynskaia, L., Wang, Q. (2004). The Experience of Rising Inequality in Russia and China During the Transition. *European Journal of Comparative Economics*, 1, 187-106.
- Goh, C.C., Luo X., Zhu, N. (2009). Income Growth, Inequality and Poverty Reduction: A Case Study of Eight Provinces in China. *China Economic Review*, 20, 485-496.
- Gustafsson, B., Sai, D. (2009). Rank, Income and Income Inequality in Urban China. *China Economic Review*, 20, 497-507.
- Hannum, E., Wang, M. (2006). Geography and Educational Inequality in China. *China Economic Review*, 17, 253-265.
- Heckman, J.J. (2005). China’s Human Capital Investment. *China Economic Review*, 16, 50-70.
- Heckman, J.J., Li, X. (2003). *Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000*. IFAU - Institute for Labour Market Policy Evaluation, Working Paper 2003,17-23.
- Huang, JRT., Kuo, C., Kao, A. (2003). The Inequality of Regional Economic Development in China Between 1991 and 2001. *Journal of Chinese Economic and Business Studies*, 1(3), 273-285.

- Jones, A.M., Nicolás, A.L. (2004). Measurement and Explanation of Socioeconomic Inequality in Health with Longitudinal Data. *Health Economics*, 13(10), 1015-1030.
- Koolman, X., Van Doorslaer, E. (2004). On the Interpretation of a Concentration Index of Inequality. *Health Economics*, 13, 649–56.
- Lahiri, K., Pulungan, Z. (2007). *Income-Related Health Disparity and Its Determinants in New York State: Racial/Ethnic and Geographical Comparisons*. MPRA Paper, 21694, University Library of Munich, Germany.
- Lee, J. (2000). Changes in the Source of China's Regional Inequality. *China Economic Review*, 11, 232-245.
- Li, H. (2003). Economic Transition and Returns to Education in China. *Economics of Education Review*, 22, 317-328.
- Li, H., Zhu, Y. (2006). Income, Income Inequality, and Health: Evidence from China. *Journal of Comparative Economics*, 34(4), 668-693.
- Liu, Z. (2005). Institution and Inequality: the *Hukou* System in China. *Journal of Comparative Economics*, 33(1), 133-157.
- Milanovic, B. (1997). A Simple Way to Calculate the Gini Coefficient, and Some Implications. *Economics Letters*, 56, 45–49.
- Mok, K.O. (2000). Marketizing Higher Education in Post-Mao China. *International Journal of Educational Development*, 20(2), 109-126.
- Ning, G. (2010). *Can Educational Expansion Improve Income Inequality in China? Evidences from the CHNS 1997 and 2006 Data*. IZA Discussion Paper, 5148.

- O'Donnell, O., Van Doorslaer, E., Wagstaff, A., Lindelöw, M. (2008). *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation*. Washington D.C.: The World Bank.
- Sicular, T., Ximing, Y., Gustafsson, B., Shi, L. (2007). The Urban-Rural Income Gap and Inequality in China. *Review of Income and Wealth*, 53(1), 93-126.
- Valli, V., Saccone, D. (2009). Structural Change and Economic Development in China and India. *European Journal of Comparative Economics*, 6(1), 101-129.
- Wagstaff, A. (2005). *Decomposing Changes in Income Inequality into Vertical and Horizontal Redistribution and Reranking, with Applications to China and Vietnam*. World Bank Policy Research Working Paper, 3559, Washington (D.C.).
- Wagstaff, A., Paci, P., Van Doorslaer, E. (1991). On the Measurement of Inequalities in Health. *Social Science & Medicine*, 33, 545–557.
- Wagstaff, A., Van Doorslaer, E., Watanabe, N. (2001). *On Decomposing the Causes of Health Sector Inequalities With an Application to Malnutrition Inequalities in Vietnam*. Policy Research Working Paper Series, 2714, The World Bank.
- Wu, X., Perloff, J.M. (2005). China's Income Distribution, 1985-2000. *The Review of Economics and Statistics*, 87(4), 763-775.
- Wu, X., Treiman, D.J. (2004). The Household Registration System and Social Stratification in China: 1955-1996. *Demography*, 41(2), 363-384.
- Wu, X., Xie, Y. (2002). *Does the Market Pay Off? Earnings Inequality and Returns to Education in Urban China*. William Davidson Working Paper, 454.
- Xing, C., Li, S. (2011). Residual Wage Inequality in Urban China, 1995–2007. *China Economic Review*, first published on-line, doi:10.1016/j.chieco.2011.10.003.

- Xu, L.C., Zou, H. (2000). Explaining the Changes of Income Distribution in China. *China Economic Review*, 11, 149-170.
- Yang, D.T. (1999). Urban-Biased Policies and Rising Income Inequality in China. *American Economic Review*, 89(2), 306-310.
- Yao, S., Zhang, Z., Hanmer, L. (2004). Growing Inequality and Poverty in China. *China Economic Review*, 15, 145-163.
- Zhang, T., Gao, S. (2003). *Matching Gender Disparity in China: Achievement and Lessons*. Paper prepared for the Education for All Global Monitoring Report 2003/4 “Gender and Education for All: The Leap to Equality”, UNESCO.
- Zhang, X., Kanbur, R. (2005). Spatial Inequality in Education and Health Care in China. *China Economic Review*, 16(2), 189-204.
- Zhang, Y., Wan, G. (2006). The Impact of Growth and Inequality on Rural Poverty in China. *Journal of Comparative Economics*, 34(4), 694-712.

# Tables

Table 1: Results of OLS regressions. Dependent variable: individual income (1000 yuan).

	1989	1997	2006
Age	0.07209**** (5.39)	0.20197**** (10.02)	0.16708** (2.22)
Age <sup>2</sup>	-0.00083**** (-5.33)	-0.00222**** (-9.75)	-0.00141* (-1.86)
Urban	-0.04210 (-0.45)	0.68264**** (4.66)	3.14392**** (6.79)
Male	0.24806**** (3.23)	0.65691**** (5.80)	2.33194**** (6.24)
Education	0.00728 (0.63)	0.10573**** (6.06)	0.68066**** (12.85)
% agriculture	0.00001 (0.13)	-0.00613** (-2.26)	-0.00718 (-1.06)
% business	0.01802**** (9.38)	0.03241**** (10.52)	0.01705* (1.91)
% wage	0.00629**** (7.08)	0.01394**** (5.17)	0.01221* (1.75)
Jiangsu	0.13857 (-1.54)	1.62538**** (7.97)	2.83156**** (-0.03)
Shandong	0.01405 (-1.94)	1.12252**** (5.19)	1.25566* (-0.81)
Henan	-0.14055 (-2.50)	-0.31749 (-1.60)	-1.76168*** (-2.29)
Hubei	-0.19833 (-2.68)	0.44120** (2.18)	-0.17119 (-1.49)
Hunan	0.41453**** (-0.60)	1.79394**** (8.07)	0.59328 (-1.13)
Guangxi	0.23833* (-1.22)	0.46671*** (2.36)	-2.21250**** (-2.52)
Constant	-0.59443** (-2.04)	-2.25140**** (-4.52)	-2.90561 (-1.43)
Income concentration index	0.0874	0.1520	0.2400
Obs.	6954	6343	5176
R <sup>2</sup>	0.035	0.191	0.130

Prob > F	0.00	0.00	0.00
----------	------	------	------

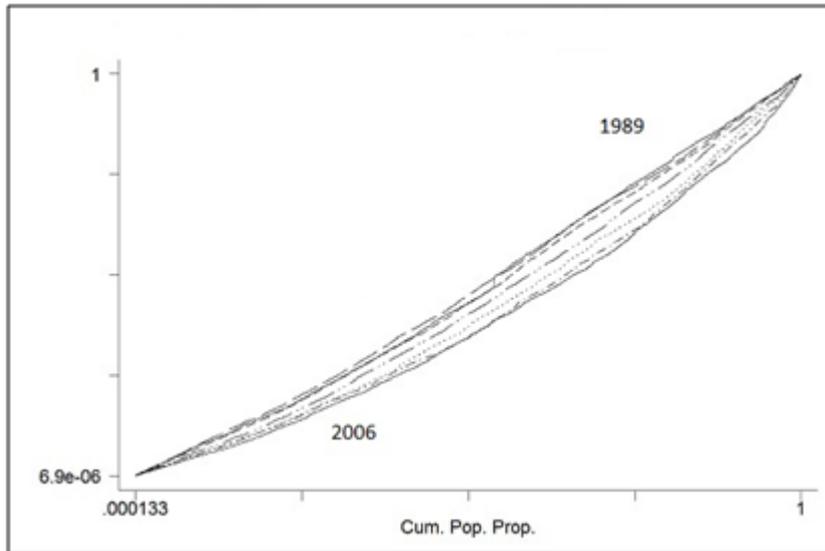
Notes: Values of t-statistics in brackets. \*, \*\*, \*\*\* and \*\*\*\* mean coefficients are significant respectively at 90%, 95%, 98% and 99.9%.

Table 2: Results of income concentration decomposition.

	1989			1997			2006		
	elasticity	CI	contrib.	elasticity	CI	contrib.	elasticity	CI	contrib.
Age	2.106	-0.099		2.014	-0.091		0.876	-0.064	
Age <sup>2</sup>	-1.064	-0.194	-2.8%	-1.049	-0.177	1.4%(total)	-0.392	-0.122	-3.6%(total)
			(total)						
Urban	-0.010	0.203	-2.3%	0.044	0.229	6.6%	0.099	0.303	12.5%
Male	0.096	0.126	13.8%	0.079	0.130	6.8%	0.126	0.117	6.2%
Education	0.030	0.422	14.5%	0.160	0.356	37.6%	0.546	0.327	74.3%
% agriculture	0.000	0.097	0.0%	-0.070	-0.224	10.3%	-0.029	-0.296	3.6%
% business	0.064	0.089	6.6%	0.089	0.099	5.8%	0.019	0.077	0.6%
% wage	0.177	0.324	65.2%	0.107	0.360	25.4%	0.049	0.273	5.5%
total			71.8%			41.5%			9.7%
Jiangsu	0.015	-0.006		0.058	0.043		0.052	0.014	
Shandong	0.001	0.005		0.032	0.003		0.018	0.046	
Henan	-0.015	-0.010		-0.012	0.019		-0.022	0.047	
Hubei	-0.022	0.008		0.016	0.011		-0.003	-0.022	
Hunan	0.041	0.088		0.046	0.153		0.008	0.143	
Guangxi	0.030	0.024	4.8% (total)	0.019	0.011	6.4%(total)	-0.036	-0.009	0.8%(total)
Residuals			0.22%			-0.21%			0.02%
Income CI		0.0876			0.1517			0.2401	

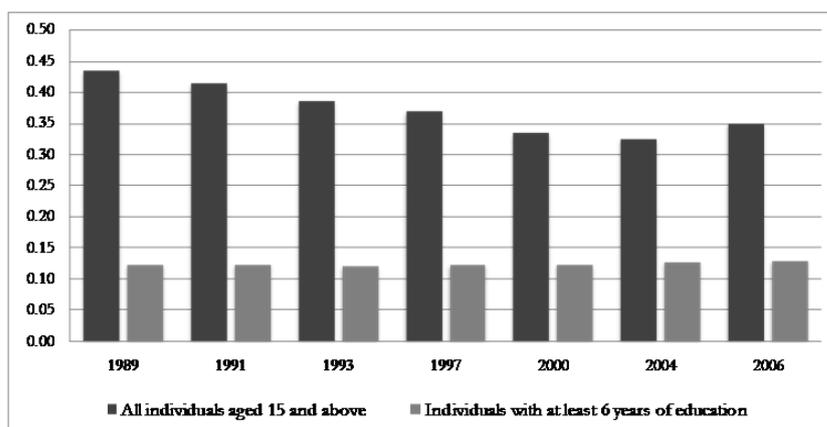
## Figures

Figure 1: Income concentration curves, 1989-2006.



Source: our calculation based on CHNS (various years). The x-axis represents the cumulative percentage of population ranked from the least to the most educated individual. The y-axis plots the cumulative percentage of individual income.

Figure 2: Gini coefficient of education, 1989-2006.



Source: our calculation based on CHNS (various years).