Increasing Returns to Education, Changing Labor Force Structure, and the Rise of Earnings Inequality in Urban China, 1996-2010*

XIANG ZHOU

University of Michigan

Word Count: (Including text, endnotes, and references: 9,922)

Number of Figures: 3

Number of Tables: 6

Number of Supplementary Files: 1

* Direct all correspondence to Xiang Zhou, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48104, USA; email: xiangzh@umich.edu. An earlier version of this paper was presented at the 2013 Annual Meeting of the American Sociological Association, and the 2013 Conference of the North American Chinese Sociologists Association. I thank Yu Xie for making the CGSS data available to me. I am grateful to Siwei Cheng, David Harding, Jeffrey Morenoff, Jeffrey Smith, Geoffrey Wodtke, Yu Xie, Hua Ye, and two anonymous reviewers for their helpful comments. The ideas expressed herein are those of the author.

Increasing Returns to Education, Changing Labor Force Structure, and the Rise of Earnings Inequality in Urban China, 1996-2010

Abstract

Earnings inequality in urban China has experienced a rapid growth over the past two decades. To account for the rise of inequality in urban China, previous studies have provided three major explanations: widening regional disparities, increasing returns to education, and growing residual inequality. Since the mid-1990s, however, the composition of the urban labor force has been dramatically altered by three large-scale structural changes: (1) the expansion of tertiary education, (2) the decline of state sector employment, and (3) a surge in rural-to-urban migration. In this article, I examine how these institutional and demographic shifts have shaped the recent upswing in earnings inequality. Based on data from two nationally representative surveys, I use variance function regressions to decompose the growth in earnings inequality from 1996 to 2010 into four components: changes in between-group earnings gaps, changes in within-group earnings variation, and two types of composition effects (*distribution effect* and *allocation effect*). I also employ counterfactual simulations to evaluate the utility of different explanations. Results show that nearly half of the growth in earnings inequality during this period is due to increases in returns to education, while the other half can be attributed to compositional changes in the labor force. The composition effects, moreover, stem chiefly from the expansion of tertiary education and the shrinkage of state sector employment.

Keywords

trends in earnings inequality, labor markets, composition effects, educational expansion, state sector shrinkage, urban China

Since the initiation of market-oriented reforms in 1978, China has experienced not only unprecedented economic growth, but also a tremendous increase in economic inequality. In 1980, the Gini coefficient for family income in China was around 0.3 (UNU-WIDER 2008), but by 2012 it was reported to have reached an alarming level of 0.61 (Gan et al. 2014; Hvistendahl 2013), a magnitude that places China among the most unequal societies in the world. While it is widely recognized that economic inequality in China is marked by a large rural-urban gap in industrial development (Knight and Song 1999; Sicular et al. 2007; Yang and Zhou 1999), recent survey data provide strong evidence that inequality within urban areas has also widened considerably over the past two decades (Jansen and Wu 2012; Li, Sato, and Sicular 2013). As shown in Figure 1, the Gini coefficient for individual earnings climbed from 0.40 in 1996 to 0.49 in 2010. The pace of this growth is striking when we consider that it took 27 years for the corresponding measure in the U.S. to increase by the same proportion: from 0.33 in 1979 to 0.41 in 2006 (McCall and Percheski 2010).

[Figure 1 here]

What are the sources of the rising inequality in urban China? How has the change in aggregate inequality been driven by changes in individual and contextual determinants of earnings? Previous research has provided three major explanations: (1) widening regional disparities (e.g., Hauser and Xie 2005), (2) increasing returns to education (e.g., Jansen and Wu 2012; Zhao and Zhou 2002), and (3) growing residual inequality (e.g., Hauser and Xie 2005; Meng, Shen, and Xue 2013). Few studies, however, have explicitly considered the role of changing labor force structure in the evolution of earnings inequality in China. In fact, since the mid-1990s, the composition of the urban labor force has been dramatically altered by three large-scale structural changes: (1) the expansion of tertiary education, (2) the decline of state sector employment, and (3) a surge in rural-to-urban migration. The main goal of this article is to examine whether, to what extent, and in what ways these institutional and demographic shifts have shaped the recent upswing of earnings inequality in urban China.

To accomplish this goal, I capitalize on variance function regressions (Western and Bloome 2009) to decompose the change in earnings inequality from 1996 to 2010 into four components: changes in between-group earnings gaps, changes in within-group earnings variation, and two types of composition effects. I also use counterfactual simulations to adjudicate between the competing explanations for the rise of inequality. Results show that nearly half of the growth in earnings inequality during this period is due to increases in returns to education, while the other half can be attributed to compositional changes in the labor force. The composition effects, moreover, are largely driven by changes in educational distribution and in sectoral structure, which in turn result from the expansion of tertiary education and the shrinkage of state sector employment.

Although focused on the context of urban China, this study sheds light on the evolution of earning inequality both in other developing countries and in other post-socialist states. On the one hand, there is a sizable body of research———in both sociology and economics———on the linkage between educational distribution and aggregate inequality in earnings (e.g., Jacobs 1985; Knight and Sabot 1983; Lam and Levison 1992; Nielsen and Alderson 1997). Would a college expansion in a developing country necessarily reduce the level of inequality? Researchers have concurred that an increase in the supply of highly educated workers can actually drive up aggregate inequality through a more dispersed educational distribution, unless this effect is offset by a drop in returns to education. My analyses lend empirical support to this proposition by showing a nontrivial contribution of college expansion to the rise of inequality in urban China. On the other hand, like China, the post-socialist countries of Central and Eastern Europe have also experienced a rapid decline in the size of state sector due to various forms of privatization, which has also been related to observed increases in economic inequality. For example, based on cross-national comparisons, Bandelj and Mahutga (2010) report a positive effect of the degree of privatization on the level of income inequality in CEE. By analyzing trends from micro-level data, the present study not only establishes this link in China, but, as we will see, also gauges the impact of state sector downsizing on earnings inequality over the past one and a half decades.

Existing Explanations

In the course of China's market-oriented transition, the rise of earnings inequality can be propelled by a wide array of social, economic, and demographic processes. Here I review three mechanisms that have been extensively discussed in the literature: widening regional disparities, increasing returns to education, and growing residual inequality.

Widening Regional Disparities

One unique feature of social inequality in China is its vast regional gaps in economic development. Back in the Mao era, different regions already varied greatly in their pace of industrialization (Kanbur and Zhang 2005). During earlier years of the market-oriented reform, regional inequality slightly narrowed; yet it widened again over the 1990s, mainly due to a persistent gap in growth rates between the coastal and the inland provinces (Wan 2007). In fact, at the beginning of the economic reform, a number of coastal cities (known as Special Economic Zones) were granted preferential policies, such as tax breaks and duty exemptions, to attract both domestic and foreign investments. Thanks to these policies, coastal provinces such as Guangdong immediately enjoyed rapid growth in both FDI and exports. These initial benefits, combined with economies of scale, soon translated into cumulative advantages (Démurger et al. 2002; Golley 2002). As a result, the coastal provinces sustained higher growth rates than the inland provinces for a long time, leading to an ever-increasing coastal-inland divide. Inequality in economic development led to differentiation in personal earnings. As shown by Xie and Hannum (1996), back in 1988, the most influential predictor of earned income in urban China was not individual attributes, but regional indicators. In a follow-up study, Hauser and Xie (2005) discover that the influence of regional differences on earnings determination increased from 1988 to 1995. While more recent trends remain unclear, there is strong evidence that regional disparities persisted, if not widened, into the 2000s. Using the 1% population sample survey of 2005, Zhang and Wu (2010) find that 41% of the total variation in earnings can be explained by between-county differences.

To the extent that there is an increase in regional inequality during the period under investigation, I aim to identify how much of the observed rise in earnings dispersion can be attributed to increased regional gaps. To accomplish this, I base my counterfactual analyses on multiple regressions that control for educational attainment and other individual attributes. This procedure helps eliminate the influence of potential confounding factors, such as increasing returns to education, a process that would exacerbate regional inequality if human capital was distributed unevenly across regions.

Increasing Returns to Education

The second potential source of rising inequality is increasing returns to education. For earlier years of China's economic reform, returns to schooling have been found to be extremely low, which is largely attributed to the absence of markets (Peng 1992; Walder 1990; Whyte and Parish 1985; Xie and Hannum 1996; Zhao and Zhou 2002). Nonetheless, the gradual expansion of markets has led theorists to predict an increase in the importance of human capital in the long term (Cao and Nee 2000; Nee 1989, 1991, 1996). This prediction has been widely supported by subsequent empirical research (Bian and Logan 1996; Hauser and Xie 2005; Wu and Xie 2003; Zhou 2000). For instance, Hauser and Xie (2005) find that in urban China, "net returns to schooling almost doubled for both men and women" from 1988 to 1995. Jansen and Wu (2012) also demonstrate a steady increase in returns to schooling over the reform period: "one additional year of schooling translated into a 2 percent net increase in income in 1978, 3.5 percent in 1985, 4.5 percent in 1990, 5.5 percent in 1995, 6.6 percent in 2000, and 7.7 percent in 2005." Yet in 1999, the Chinese government implemented a college expansion policy that suddenly increased college enrollments by almost a half. This policy significantly raised the supply of college-educated workers over the following decade, which may have slowed down the growth in returns to education (Meng et al. 2013).

How would an increase in returns to education influence the size of earnings inequality? Xie and Hannum (1996) show that, holding constant the marginal distribution of human capital, an increase in returns to schooling would generally drive up the overall inequality. Thus, I expect the rise of inequality during the study period to be partly explained by an increase in returns to education, although the size of this increase since the early 2000s may have been moderated by an expanding supply of college-educated workers. As with changing regional gaps, the impact of changing returns to education on earnings inequality will be assessed by counterfactual analyses.

Growing Residual Inequality

Beyond changes in observed determinants of earnings, a third explanation for the rise of earnings inequality is growing residual variation. Labor economists studying inequality in the U.S. have found that the rise of wage inequality in the 1970s and 1980s is primarily due to an increased residual variance of earnings after individual-level predictors such as schooling, experience, and demographic attributes are factored in (Juhn, Murphy, and Pierce 1993). This finding has been closely linked to the theory of "skill-biased technological change" (henceforth SBTC), which posits that growing residual inequality is mainly a result of rising returns to unobserved skills among workers with the same observed characteristics (Acemoglu 2002). Similar to trends in the U.S., the rise of earnings inequality in urban China from the late 1980s to the mid-1990s has also been related to an increase in residual variation (Hauser and Xie 2005).

While traditional regression-based analyses assume homoscedasticity and thus regard residual variance as uniform among all individuals, recent literature on inequality has begun to address heterogeneity in residual variance across population subgroups (Lemieux 2006; Western and Bloome 2009). When this heterogeneity is taken into account, the change in total residual inequality between two time points consists of two components: one represents changes in residual inequality among people in the same observed groups, and the other reflects the effect of changing group proportions. In fact, Lemieux (2006) challenges the SBTC explanation by showing that the growth of residual inequality in the U.S. during the1990s was mainly propelled by changes in the proportion of workers in different experience-education cells, rather than changes in within-cell variation. In this study, I also separate out these two drivers of residual inequality by modeling sectoral differences in residual variation in China. Specifically, I consider the first component as an essential change in residual inequality, and use *allocation effect* to mean the impact on residual inequality of changes in labor force composition. For example, if inequality is greater in the private sector than in the state sector, a shift in the workforce from the state sector to the private sector can produce an allocation effect that amplifies the level of overall inequality.

A Missing Link: Composition Effects

Among the above explanations, widening regional disparities and increasing returns to education can be construed as changing earnings gaps between population subgroups (in these cases, region and education), whereas growing residual inequality reflects increases in within-group variation. If the composition of the labor force is fixed, all sources of change in overall inequality can be subsumed under these two categories. Nonetheless, when group proportions are time-varying, trends in aggregate inequality can also be driven by composition effects. In fact, since the mid-1990s, the composition of the labor force in urban China has been dramatically reshaped by three large-scale socio-economic changes: (1) the expansion of tertiary education, (2) the decline of state sector employment, and (3) a surge in rural-to-urban migration (for more details, see Figure S1 in supplementary data). In this section, I discuss how these compositional shifts may have contributed to the rise of earnings inequality during the past two decades.

Expansion of Tertiary Education

As noted above, in 1999, the Chinese government adopted a collage expansion policy that significantly enlarged the pool of college-educated workers over the following decade. The purpose of this policy was two-fold. First, it was aimed to increase the supply of skilled labor for sustaining China's rapid economic growth. Second, extending schooling for the youth was considered as a strategy to alleviate the pressure of re-employment for those who were being laid off during the reform of state-owned enterprises (see the next subsection). Coupled with cohort replacement, since 2003, the expansion of higher education has triggered a rapid change in the distribution of human capital among the urban labor force. In 2003, those who had finished at least a three-year college constituted only 9.1% of the urban population (aged 6+); but by 2010, this portion had more than doubled to 21.5% (see Figure S1).

What is the implication of such a compositional shift for earnings inequality? In fact, before China's expansion of tertiary education, the educational distribution among urban workers was highly concentrated at the levels of junior and senior high school, suggesting a relatively homogeneous labor force in terms of observed skills. However, as more youth were provided the opportunity of obtaining a college degree, cohort replacement resulted in a more dispersed educational distribution, which, everything else being equal, should have amplified earnings inequality in the aggregate. Thus, I expect that *the rise of earnings inequality in urban China can be partly attributed to changes in educational distribution.*

Shrinkage of State Sector Employment

As with other post-socialist countries, one central aspect of China's economic transition is the decline of state sector employment. Although the market-oriented reform in urban China started as early as 1984, it was concentrated on the goods market during its first decade. In the early 1990s, the vast majority of urban workers were still employed in state-owned enterprises (henceforth SOE), the prototypical work unit in pre-reform urban China. By 1994, however, most of the SOEs had excessive employment and nearly half of them were making losses, severely hindering China's economic development (Cao, Qian, and Weingast 2003). To remedy this situation, beginning in 1995, the Chinese government has been reforming and downsizing state-owned enterprises under the policy of "grasp the large and let go the small." On the one hand, the central government began to merge and restructure large SOEs, thereby consolidating its control over certain strategically vital industries, such as power generation, telecommunication, and raw materials. On the other hand, at the local level, small SOEs were largely privatized, and workers in medium-size SOEs were laid off massively. As a result, since the mid-1990s, tens of millions of former SOE employees have been pushed into the private sector. Among new entrants to the labor market, the share of state sector employment has also dwindled. Such an imbalance between exit and entry has produced a steady decline in state sector employment during the past two decades: in 1996, 64% of the urban workers were employed in the state sector, but by 2010 this figure had reduced to 27% (see Figure S1).

It is widely acknowledged that the SOE reform has been successful in vitalizing China's market economy. In the meantime, however, the massive transfer of labor from the state sector to the private sector may have exacerbated the country's earnings distribution. Before the SOE reform, the majority of urban workers were employed by the state with a centrally-planned wage system, which imposed a highly compressed earnings distribution. Earnings variation within the state sector was mainly driven by differences in bonus income, which depended heavily on the profitability of their work units (Wu 2002; Xie and Wu 2008). Overall, earnings inequality was substantially lower in the state sector than in the private sector, partly because observed and unobserved skills were less rewarded by the state, and partly because the paychecks of state employees were less sensitive to the ebb and flow of the market. This pattern, in fact, has been fairly stable over time. Today's SOEs in China continue to benefit from sheltered markets, implicit government subsidies, and politically favored bank loans. By shielding the SOEs from market competition, these institutional protections have sustained a relatively low dispersion of earnings across the state sector. Meanwhile, the downsizing of SOEs has pushed tens of millions of workers into the private sector, where their heterogeneity in ability and skills is more likely to translate into different rates of pay. Therefore, given that earnings inequality is lower in the state sector than in the private sector, I hypothesize that *the massive transfer of workers from the former to the latter has contributed to the observed rise in aggregate inequality.*

Rural-to-urban Migration

In the pre-reform era, rural-urban migration in China was severely restricted by the Chinese household registration system, i.e., *hukou*, a state institution established to limit population mobility. Since 1978, the market-oriented reform has moderately eased the restriction for temporary migration, but not for conversion of *hukou* status, thus resulting in a "floating population" of urban dwellers with rural *hukou* status (Wu and Treiman 2004). The size of this floating population was relatively small, if not negligible, until the early 1990s. Since then, China's economic growth has been increasingly propelled by export-oriented manufacturing sectors and government-sponsored infrastructure projects, which have significantly raised the demand for young and low-skilled workers in many urban centers. The surge of demand for cheap labor has attracted wave after wave of young and poorly-educated migrants from the rural inland. As a result, the volume of rural migrants residing in urban centers has increased tremendously over the past two decades. According to Meng et al. (2013), the number of rural-urban migrant workers was about 39 million in 1997, and by 2009, the size had increased to 145 million, constituting more than a quarter of the urban labor force.

In spite of their growing contribution to the economic boom in urban areas, it remains extremely difficult for these rural migrants to acquire a local *hukou* in the cities where they work. As noted by Chan and Buckingham (2008), in such large cities as Beijing, Shanghai and Guangzhou, which are the major destinations of recent waves of rural-urban migrants, the entry requirements for obtaining a local *hukou* are highly prohibitive and clearly beyond the reach of most migrant workers. The lack of local *hukou* status is perhaps the greatest curse for this ever-increasing floating population, because *hukou* status was and still is a very strong institutional constraint shaping one's social and economic wellbeing in urban China (Treiman 2012; Wu and Treiman 2004, 2007). Not only is local *hukou* status a prerequisite for such social welfare benefits as health care and unemployment insurance, but migrant workers without a local *hukou* also suffer from a variety of unfair treatment in the workplace, such as wage arrears and denial of payments.

Given the persistent power of *hukou* in affecting one's economic wellbeing, how may the recent upsurge in rural-to-urban migration have affected earnings inequality in urban China? Indeed, as shown by Meng and Zhang (2001), in the 1990s, migrant workers without an urban *hukou* were subject to a wage penalty in the urban labor market. Yet it is unclear whether such a wage gap narrowed or widened into the 2000s, and whether the wage gap necessarily translated into an earnings gap between the two groups (considering that migrant workers typically work for longer hours and more days than local urban workers). However, *to the extent that an earnings gap exists across the hukou axis, the surge in rural-to-urban migration should have subjected a larger share of the workforce to an earnings penalty, thereby triggering an increase in overall inequality.*

Methods

R^2-based Methods

In this study, I use the variance of log earnings to gauge the size of earnings inequality. The variance measure is particularly useful for studying trends in inequality because it can be easily decomposed into between-group and within-group components using ANOVA (see Mouw and Kalleberg 2010). The ratio of the between-group component to the total variance provides an intuitive measure for the between-group contribution to total inequality, which would be equivalently given by the R^2 in a linear regression of log earnings on group dummies. To examine temporal trends in the size of between-group contribution, one may simply track changes in this ratio over time. For example, Kim and Sakamoto (2008) used the time series of occupation R^2 to assess the relative importance of between-occupation and within-occupation inequality in explaining the rise of wage inequality in the U.S. Moreover, in a regression model that controls for covariates, we can evaluate the net contribution of a particular set of variables using incremental or partial R^2s (see Kim and Sakamoto 2008; Meng et al. 2013). As a preliminary analysis, I also use partial R^2 to detect variations in the importance of different earnings determinants over time.

This approach, however, is prone to conflate changes in population composition with *real* changes in between-group disparities and within-group variation. To see this, consider a hypothetical population consisting of two groups: college graduates and high school graduates. Assume that the average gap in log earnings between the two groups is fixed, and that the within-group variation among college graduates is greater than that among high school graduates. Now consider an education expansion that enlarges the share of college graduates from 10% to 50%. In this case, earnings inequality will increase, neither via increased returns to education nor via increased within-group inequality, but via a change in population composition. Specifically, the impacts of this compositional shift are two-fold. On the one hand, given an earnings premium for college graduates, a more balanced distribution of the two groups will automatically inflate the overall variance. On the other hand, given that within-group inequality is higher among college graduates than among high school graduates, an increased share of the former will also raise the level of total inequality. The R^2 measure, however, can drift in either direction without a clear interpretation.

Variance Function Regressions and Decomposing Trends in Inequality

My analytical focus is to disentangle different sources of the observed rise in earnings inequality, thus adjudicating between the competing explanations discussed in the preceding sections. To achieve this goal, I decompose the change in the variance of log earnings based on variance function regressions (Western and Bloome 2009), a technique that allows both the mean and the variance of log earnings to depend on a set of explanatory variables.

To sketch this approach, let us denote by Y_t the dependent variable, log earnings, at time t. Meanwhile, denote by X_t and Z_t two sets of independent variables that are used to predict the mean and the variance of log earnings, respectively. We then jointly estimate the conditional mean and conditional variance of log earnings respectively as linear functions of X_t and log-linear functions of Z_t, yielding two fitted models:

$$\hat{E}(Y_t|X_t) = \hat{\beta}_t X_t, \qquad \widehat{Var}(Y_t|Z_t) = exp\ (\hat{\lambda}_t Z_t),$$

where $\hat{\beta}_t$ and $\hat{\lambda}_t$ denote estimated coefficients of X_t and Z_t. As a result, the fitted total variance of log earnings can be written as

$$\hat{V}_t = \widehat{Var}\big[\hat{E}(Y_t|X_t)\big] + \hat{E}\big[\widehat{Var}(Y_t|Z_t)\big] = \widehat{Var}\big(\hat{\beta}_t X_t\big) + \hat{E}\big[exp\big(\hat{\lambda}_t Z_t\big)\big]. \tag{1}$$

This equation can be seen as a parametric analog of ANOVA, with the first term and the second term corresponding respectively to between-group and within-group components of inequality. According to equation (1), the change in earnings inequality from time t to another time point t' ($t < t'$) can be written as

$$\hat{V}_{t'} - \hat{V}_t = \widehat{Var}\big(\hat{\beta}_{t'} X_{t'}\big) - \widehat{Var}\big(\hat{\beta}_t X_t\big) + \hat{E}\big[exp\big(\hat{\lambda}_{t'} Z_{t'}\big)\big] - \hat{E}\big[exp\big(\hat{\lambda}_t Z_t\big)\big], \tag{2}$$

where the first contrast $\widehat{Var}\big(\hat{\beta}_{t'} X_{t'}\big) - \widehat{Var}\big(\hat{\beta}_t X_t\big)$ measures the change in between-group inequality, and the second contrast $\hat{E}\big[exp\big(\hat{\lambda}_{t'} Z_{t'}\big)\big] - \hat{E}\big[exp\big(\hat{\lambda}_t Z_t\big)\big]$ gauges the change in within-group inequality. These two parts can be further decomposed to separate the effects of changing coefficients (β and λ) from those of changing distributions of X and Z. Specifically, equation (2) can be expanded as

$$\hat{V}_{t'} - \hat{V}_t = \hat{\delta}_B + \hat{\delta}_D + \hat{\delta}_W + \hat{\delta}_A, \tag{3}$$

with

$$\hat{\delta}_B = \widehat{Var}\big(\hat{\beta}_{t'} X_t\big) - \widehat{Var}\big(\hat{\beta}_t X_t\big)$$

$$\hat{\delta}_D = \widehat{Var}\big(\hat{\beta}_{t'} X_{t'}\big) - \widehat{Var}\big(\hat{\beta}_{t'} X_t\big)$$

$$\hat{\delta}_W = \hat{E}\left[exp\left(\hat{\lambda}_{t'} Z_t\right)\right] - \hat{E}\left[exp\left(\hat{\lambda}_t Z_t\right)\right]$$

$$\hat{\delta}_A = \hat{E}\left[exp\left(\hat{\lambda}_{t'} Z_{t'}\right)\right] - \hat{E}\left[exp\left(\hat{\lambda}_{t'} Z_t\right)\right].$$

In this decomposition, the first term, δ_B, measures the change in between-group earnings gaps. For example, if region is the only predictor of earnings, then δ_B represents the impact of widening (if $\delta_B > 0$) or narrowing (if $\delta_B < 0$) regional disparities on total inequality. The second term, δ_D, gauges the change in between-group inequality due to changes in population composition. Recent research on the U.S. labor market has revealed a polarization of the occupational structure, i.e., growing employment in both high- and low- paying occupations and hollowing out of the middle (Massey and Hirst 1998; Mouw and Kalleberg 2010). Such compositional changes would drive up overall inequality even if between-occupation differences in average earnings were fixed. For this reason, I refer to δ_D as *distribution effect*. Clearly, changes in between-group gaps (δ_B) and the distribution effect (δ_D) together constitute the total change in between-group inequality ($\delta_B + \delta_D$). The third term, δ_W, characterizes the change in within-group variation among people with the same observed characteristics. In the economics literature, this component has an intimate connection with the theory of SBTC, which stresses the role of increasing returns to skills (often unobserved) in the growth of residual inequality. The last term, δ_A, identifies the change in within-group inequality due to changes in population composition. As discussed in the preceding section, during the SOE reform, the massive labor transfer from the state sector to the private sector may have induced an increase in overall inequality as a result of unequal residual variations between the two sectors, even if the amounts of within-sector inequality stayed unchanged over time. Hence, I term δ_A *allocation effect*. The separation of the allocation effect from δ_W enables us to disentangle the impacts of compositional shifts in the labor force from more inherent changes in residual inequality. The structure of this four-component decomposition is shown more lucidly in Table 1.

[Table 1 here]

Note that the above decomposition is not mathematically unique. In equation (3), the difference between $V_{t'}$ and V_t is decomposed in a way that changes in coefficients happen first and changes in population composition come second. Reserving this order would yield an alternative decomposition. Below, I use Type I decomposition to mean equation (3) and call the alternative one Type II decomposition.

Counterfactual Analysis

Results from variance function regressions can be used to construct counterfactual levels of inequality, thus enabling us to assess the utility of competing explanations (Western and Bloome 2009).[1] For example, to evaluate the effect of changing returns to education, we can calculate the following counterfactual:

$$\hat{V}_{t'}^{\beta^{edu}=\hat{\beta}_t^{edu}} = \widehat{Var}\left(\hat{\beta}_{t'}^{-edu} X_{t'}^{-edu} + \hat{\beta}_t^{edu} X_{t'}^{edu}\right) + \hat{E}[exp\ (\hat{\lambda}_{t'} Z_{t'})], \tag{4}$$

where β^{edu} denotes the coefficient (or a set of coefficients) for education, and β^{-edu} denotes the coefficients for all other predictors. Equation (4) gauges the level of inequality that would have been observed at time t' had returns to education stayed at the level of time t. Thus, the difference between $\hat{V}_{t'}$ and $\hat{V}_{t'}^{\beta^{edu}=\beta_t^{edu}}$ identifies the contribution of changing returns to education to the change in overall inequality from t to t'.

To assess the impact of a compositional shift, we can reweight the observed data at time t' to make the marginal distribution of the corresponding variable identical to that at time t (see Lemieux 2006). For instance, to examine the influence of college expansion on earnings inequality, we can fix the marginal distribution of educational attainment at time t by appropriately down-weighting college graduates and up-weighting others in the sample at time t', i.e.,

$$\hat{V}_{t'}^{\pi^{edu}=\pi_t^{edu}} = \widehat{Var}_{\pi_t^{edu}}\left(\hat{\beta}_{t'} X_{t'}\right) + \hat{E}_{\pi_t^{edu}}[exp\ (\hat{\lambda}_{t'} Z_{t'})], \tag{5}$$

where π_t^{edu} denotes the educational distribution at time t, and its appearance as subscript means that corresponding weights are used to calculate the variance and the expectation. The composition effect due to changing educational distribution is thus measured by the difference between $\hat{V}_{t'}$ and $\hat{V}_{t'}^{\pi^{edu}=\pi_t^{edu}}$:

$$\hat{V}_{t'} - \hat{V}_{t'}^{\pi^{edu}=\pi_t^{edu}} = \widehat{Var}(\hat{\beta}_{t'}X_{t'}) - \widehat{Var}_{\pi_t^{edu}}(\hat{\beta}_{t'}X_{t'}) + \hat{E}[exp\ (\hat{\lambda}_{t'}Z_{t'})] - \hat{E}_{\pi_t^{edu}}[exp\ (\hat{\lambda}_{t'}Z_{t'})].$$

The above expression reveals that the composition effect consists of two parts, representing changes in between-group and within-group inequalities. Hence, the first part identifies the *distribution effect*, and the second part identifies the *allocation effect*.

While the above illustrations are both for the variable of education, the same techniques can be employed to gauge the effects of changes in other determinants of earnings. Table 2 shows how the competing explanations discussed earlier are to be examined by counterfactual analysis. For example, by reweighting the 2010 data such that the sectoral composition equals that in 1996, we can assess the allocation effect of state sector shrinkage. However, because the educational distribution can systematically differ across sectors, the reweighting method is unable to manipulate the marginal distribution of one variable without changing that of the other. Therefore, in the following analysis, I also examine the combined effects of changing educational and sectoral compositions by fixing their joint distributions at the 1996 level.

[Table 2 here]

Data

I use data from two nationally representative sample surveys: the 1996 survey of "Life History and Social Changes in Contemporary China" (henceforth LHSC 1996) and the 2010 wave of the Chinese General Social Survey (henceforth CGSS 2010). Although these two surveys have different names, their data are highly comparable for my trend analysis. First, both surveys adopted a multi-stage stratified sampling design under which one adult was randomly selected from each sampled household (Li and Wang 2012; Treiman and Walder 1998). Second, in both surveys, the fieldwork was implemented by the same organization——the Department of Sociology at Renmin University of China. Moreover, they used the same rule to demarcate urban and rural populations——namely, whether the sampled household belonged to a neighborhood committee (urban) or a village committee (rural)——which ensures that the two urban samples are consistent in their coverage.

While CGSS 2010 collected data from all 31 provinces of mainland China, the sampling frame of LHSC 1996 did not include Tibet. To maintain the comparability of labor markets over time, I excluded Tibet from the CGSS 2010 data as well (step 1: N_{1996}=3087, N_{2010}=7081). Since Tibet represents only 0.2% of the Chinese population (National Bureau of Statistics of China 2011), its exclusion is unlikely to weaken the representativeness of the data. To assess earnings inequality among the economically active population, I further restricted both samples to those who were between ages 20 and 69 and regularly employed with annual earnings greater than 100 1996 Yuan (step 2: N_{1996}=2024, N_{2010}=3050).[2] After eliminating a small number of respondents with missing covariates, we have 2019 and 3040 individual workers from LHSC 1996 and CGSS 2010, respectively.

In this study, the dependent variable, earnings, refers to the total amount of earned income, including wages and salaries, bonuses, and profits from private businesses.[3] Earnings in 1996 are inflation-adjusted to 2010 Yuan based on official CPI rates (National Bureau of Statistics of China 2011). To adjudicate between the competing explanations for the rise of earnings inequality, I use the following explanatory variables: province, education, sector of employment, and *hukou* status. To better address composition effects, education is treated as a categorical variable containing six levels of educational attainment: (1) no schooling, (2) elementary school, (3) junior high school, (4) senior high school or vocational high school, (5) vocational college, (6) four-year college or above. While most previous studies treated sector of employment as a state-market dichotomy, I adopt a finer typology of sector: (1) state sector, which includes government agencies, public organizations, and state-owned enterprises, (2) private sector, which includes domestic private enterprises, foreign-invested firms, joint ventures, as well as collective enterprises and institutions,[4] and (3) self-employment. *Hukou* status is coded as a binary variable (non-agricultural vs. agricultural) in order to identify rural-urban migrants. The regression model for the mean of log earnings also includes sex, age, age squared, and party membership as covariates.

Table 3 reports some descriptive statistics. The first two columns show the population share of different subgroups in 1996 and 2010. With regard to sex, age, and party membership, the group proportions are fairly similar across the two years, although the workforce appears slightly older in 2010. The share of workers holding a rural *hukou* increased sharply, from 12% in 1996 to 27% in 2010, reflecting the sheer scale of rural-to-urban migration. Thanks to college expansion, the proportion of workers who had obtained a college degree (either vocational or regular) more than doubled. Moreover, state sector employment experienced a dramatic decline: in 1996, 59% of the workers were employed in the state sector, but by 2010 this portion had reduced to 27%.

[Table 3 here]

The next two columns present the group-specific means of log earnings. Overall, we see a substantial increase in earnings for both men and women, both party members and non-members, and all age groups. However, on average, earnings growth seems to be more significant for permanent urban dwellers and more-educated workers than for rural-urban migrants and less-educated workers. The last two columns demonstrate the group-specific levels of inequality, measured by the variance of log earnings. We find that the rise of earnings inequality is greater among party members and permanent urban dwellers than among non-members and rural-urban migrants. Moreover, for both years, earnings dispersion is much lower in the state sector than in the private sector, and the self-employed exhibit the highest within-group inequality.

Results

Partial R^2s from Conventional Regressions

To gauge the influence of a given set of variables on earnings inequality, past research has often relied on R^2 or partial R^2 from multiple regressions of log earnings. As discussed earlier, this approach is not well suited for studying trends in inequality because it is prone to conflate changes in population composition with inherent changes in between-group gaps and within-group variation. For a particular time point, though, it can provide a snapshot of the structure of earnings inequality. In Figure 2, the bar plots show the net contributions of province, education, sector of employment, and *hukou* status to the overall inequality, measured by the corresponding partial R^2s. First, we find that province is the most influential factor shaping earnings inequality in urban China: in both years, nearly 15% of the variation in log earnings can be explained by interprovincial disparities, even after covariates such as sex, age, and education are controlled for. Second, we see a sharp increase in the importance of education: the partial R^2 grew from 4.7% in 1996 to 12.3% in 2010. Finally, sector of employment accounts for roughly 3% of total inequality at both time points, and the explanatory power of *hukou* status is negligible for both years.

[Figure 2 here]

The above results highlight the significance of region and education in maintaining earnings inequality in urban China. Nonetheless, they do not allow us to separate out different sources of the growth in inequality. For example, the rise in the partial R^2 of education could stem from real increases in returns to education, or changes in educational composition (i.e., distribution effect), or both. I now turn to results from variance function regressions, which provide a basis for both decomposition and counterfactual analyses.

Variance Function Regressions and Decomposition of the Rise in Inequality

Table 4 reports the results from variance function regressions. The first two columns present the effects of different predictors on the mean of log earnings. First, for both years, we see an earnings penalty for females, a premium for party-members, and a quadratic effect of age, which are all consistent with past research on earnings determination in urban China (e.g., Xie and Hannum 1996). However, we find that the effect of rural *hukou* is not significantly different from zero in either 1996 or 2010, indicating that there may not be an earnings penalty for rural-urban migrants when covariates, such as education and sector, are factored in. Meanwhile, we see a sharp increase in economic returns to a college degree (either vocational or regular) over this period: in 1996, a worker with a four-year college education was expected to earn 30% ($e^{0.264} - 1$) more than a worker with only a high school diploma; by 2010, this gap had widened to 84% ($e^{0.608} - 1$).[5] In addition, for both years, we observe an earnings premium for workers in the state sector compared with employees in the private sector. The self-employed seem to have improved their position dramatically: in 1996, they earned markedly less than the other two groups; by 2010, they had become the most advantaged group, earning about 20% ($e^{0.183} - 1$) more than state sector workers.

[Table 4 here]

My earlier argument suggests that residual inequality can be substantially lower in the state sector than in the private sector. To model sectoral differences in residual inequality, sector dummies are used as predictors in the variance regressions.[6] As shown in the last two columns, estimated residual variation is much smaller in the state sector than in the private sector, and the self-employed are the most unequal group. This pattern holds true in both years, although to a lesser extent in 2010 than in 1996. This heterogeneity in residual variance underlies my hypothesis that the decline of state sector employment can drive up the level of overall inequality via an allocation effect.

Based on the coefficient estimates in Table 4, the change in inequality from 1996 to 2010 can be decomposed into the four components expressed by equation (3). The bar plots in Figure 3 show the results from both Type I and Type II decompositions. We find that changes in between-group earnings gaps account for 34%-46% of the total growth in earnings inequality, depending on the way the decomposition is performed. Distribution effect explains 22%-34% of the total growth, whereas the contribution of allocation effect ranges from 21% to 37%. Taking them together, we see that more than half of the rise in inequality over this period can be attributed to compositional shifts in individual and contextual characteristics. By contrast, the contribution of δ_W ranges from -5%-12%, suggesting that changes in within-group dispersion have very small if any impact on the change in earnings inequality over this period.

[Figure 3 here]

Counterfactual Analyses: Evaluation of Competing Mechanisms

I now evaluate the utility of different explanations through counterfactual analyses. In Table 5, the first column presents the variances of log earnings adjusted for changes in between-group gaps (i.e., β) and in within-group variation (i.e., λ), and the second column shows the counterfactual change from 1996 to 2010 when between-group/within-group effects are fixed at the 1996 level. The third column reports the percentage of the total change explained, that is, other things being equal, how much of the total rise in inequality would have disappeared had the corresponding between-group/within-group effects stayed unchanged during this period. First, fixing the coefficients of province dummies yields an adjusted variance of 0.839, suggesting that changing interprovincial disparities accounts for none of the total growth in inequality. In contrast, by fixing the coefficients of educational attainment, we find that rising returns to education explains 45.2% of the total growth. The next row shows that if all between-group earnings gaps had stayed at the 1996 level, 45.8% of the increased inequality would have disappeared. Comparing these two numbers, we conclude that changes in between-group gaps are almost entirely driven by increases in returns to education. Finally, by fixing the coefficients in the variance model (λ), we find that changes in within-sector earnings variation have virtually no influence on the rise of inequality over this period.

[Table 5 here]

Table 6 shows the variances of log earnings adjusted for a variety of compositional shifts, together with the contributions of distribution effects, allocation effects, and total composition effects. First, we find that the distribution effect of changing *hukou* composition is close to nil, which mirrors the fact that rural *hukou* is not statistically significant for predicting log earnings. In other words, because there is not a discernible gap in earnings between rural-urban migrants and permanent urban workers, changing *hukou* composition has little impact on the trends in earnings inequality. Second, the distribution effect of education, which results chiefly from the college expansion policy, accounts for 21.9% of the total change in inequality. That is, more than a fifth of the increased variation in log earnings can be attributed to a more dispersed distribution in educational attainment.[7] Third, the allocation effect due to changes in sectoral composition also explains about one fifth of the increased inequality. This effect reflects the crucial role of state sector shrinkage: Because within-sector variation is substantially lower in the state sector than in the private sector, the massive labor influx into the latter has inflated earnings inequality in the aggregate.

[Table 6 here]

Although we do not assume any effects of *hukou* and education on the variance of log earnings, both changing *hukou* composition and changing educational composition exhibit allocation effects as well. This is because the distributions of *hukou* and of educational attainment are not independent of the distribution of sector of employment. In fact, according to the 2010 data, rural-urban migrants are more likely to work in the private sector than permanent urban workers, and college-educated workers are more likely to work in the state sector than other educational groups. Therefore, a down-weighting of rural-urban migrants would lower the average within-group inequality, whereas a down-weighting of college-educated workers would heighten it. As a result, we observe a positive allocation effect of rural-urban migration and a negative allocation effect of changing educational composition. These allocation effects, however, should not be taken at face value because the compositional shifts of *hukou* and education can be closely intertwined with changes in sectoral structure. Hence, I proceed to examine the combined effects of different compositional shifts by fixing the joint distribution of the corresponding variables. In particular, by fixing the joint distribution of education and sector at that of 1996, we find that 41.9% of the total increase in inequality results from compositional changes in education and sector of employment. This number, not surprisingly, roughly equals the sum of the distribution effect due to changing educational composition and the allocation effect due to changing sectoral composition. Finally, when the joint distribution of all observed characteristics (i.e., the data matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$) is fixed at the 1996 level, the increased variance from 1996 to 2010 drops from 0.304 to 0.137, suggesting that 54.9% of the total growth in inequality is due to compositional shifts in individual and contextual characteristics. Of these composition effects, about three quarters (41.9%/54.9%=76.3%) come from changing educational and sectoral distributions.

In short, my findings indicate that the rise of earnings inequality from 1996 to 2010 is primarily driven by (1) increases in returns to education, (2) a more dispersed educational distribution, and (3) changes in sectoral structure. In particular, the composition effects of (2) and (3) stem chiefly from the policy of college expansion and the institutional downsizing of state-owned enterprises.

Conclusion and Discussion

Earnings inequality in urban China has experienced a dramatic growth over the past two decades. To account for the rise of inequality in urban China, prior studies have offered three major explanations: widening regional gaps, increasing educational returns, and growing residual inequality. In this article, I discussed how the recent upswing in earning inequality can also be shaped by three large-scale structural changes: (1) college expansion, (2) state sector shrinkage, and (3) rural-to-urban migration. To adjudicate between existing explanations and these composition effects, I capitalized on variance function regressions to decompose and simulate the change in earnings inequality between 1996 and 2010. My results suggest that nearly half of the growth in earnings inequality during this period can be explained by increases in returns to education, while the other half is attributable to compositional shifts in the labor force. The composition effects are mainly due to changes in educational and sectoral distributions, which result respectively from the expansion of tertiary education and the shrinkage of state sector employment.

In addition, we find little effect of the upsurge in rural-urban migration on earnings inequality. In fact, my regression results show no significant difference in earnings between rural migrant workers and permanent urban workers once covariates, such as education and sector, are taken into account. This finding, however, does not contradict earlier studies that demonstrate a wage penalty for rural migrant workers (Meng and Zhang 2001), because a wage penalty does not necessarily imply a gap in earnings considering that rural migrants usually work for longer hours and more days than local urban workers. In addition, it is worth noting that although rural-urban migration seems to have limited impact on earnings inequality in urban China, it may have a profound influence on economic inequality in China as a whole. Assuming that migrant workers earn more in urban areas than they would in their rural origins, an increasing volume of migrant workers can narrow the gap between these two otherwise segregated and unequal populations (i.e., urban and rural *hukou* holders), thereby reducing the level of nationwide inequality.

Methodologically, this article illustrates the utility of variance function regressions, a technique recently proposed by Western and Bloome (2009), for studying trends in inequality.

By simultaneously modeling the mean and the variance of log earnings, this method allows the change in earnings inequality to be decomposed into four components: changes in between-group gaps (δ_B), changes in within-group variation (δ_W), distribution effect (δ_D), and allocation effect (δ_A). Different from R^2-based methods, this approach distinguishes the *dynamics* of inequality (i.e., studies on the *change* of inequality) from the *statics* of inequality (i.e., studies on the *level* of inequality). In a society, the principal factors that maintain the level of inequality do not necessarily correspond to the major forces that drive the change in inequality. For example, we find that, on the one hand, province is the most salient mediator of earnings inequality in urban China (see Figure 2), but on the other hand, the observed rise in inequality since the mid-1990s is not much driven by widening provincial disparities, but chiefly propelled by increasing returns to education and composition effects. Yet an assessment of trends in R^2 will not disentangle composition effects from inherent changes in between-group gaps or within-group variation. For instance, Figure 2 has shown a tremendous growth in the partial R^2 of education, which, however, does not necessarily result from an increase in returns to education. Without an explicit decomposition of the trend, we cannot separate the effect of changing returns to education from the effect of changing educational distribution. Similarly, without an explicit modeling of heteroscedasticity across employment sectors, we would conflate real changes in within-sector inequality with shifts in sectoral composition.

Substantively, this study provides new insights into the way economic inequality can be shaped by rapid socio-structural changes. For example, standard economic theory predicts that *ceteris paribus*, an educational expansion will cause a decline in returns to schooling due to increased market competition. By this logic, if educational expansion produces a composition effect that drives up earning inequality, it can be offset or even outweighed by a drop in returns to education. Indeed, this effect has been observed in both African and Latin American countries (Knight and Sabot 1983; World Bank 2011). This study, however, depicts a different picture for

China: since the mid-1990s, returns to higher education have increased in spite of a growing supply of college-educated workers. As a result, these two forces have operated in the same direction toward a higher level of inequality. Thus, my finding indicates that the actual effects of educational expansion on inequality cannot always be reduced to a "partial equilibrium analysis;" instead, they are simultaneously shaped by a variety of supply-side, demand-side, and non-market processes in a historical context.

While my analyses have broadly linked the growth in inequality to observed changes in earnings determinants, they are limited in revealing the complexity of micro-level processes. For example, although the observed increase in returns to education comports with the market transition theory, it is not necessarily due to market forces per se. First, if students with more (unobserved) family resources selectively obtained more education, the increase in estimated returns to education would reflect an increase in the compounded effects of schooling and family resources. Second, during the economic reform, state bureaucracies have also increasingly emphasized educational credential in resource allocation, which may have also raised the observed returns to education. In fact, due to state sponsorship, part-time adult colleges——which confer nearly a third of undergraduate diplomas in China——are much more likely to recruit mid-career cadres and state professionals than less privileged individuals (Lai 2014). If this effect had intensified over the period under investigation, the observed increase in returns to college may also have been inflated.

The results of variance regressions show a markedly lower level of inequality in the state sector than in the private sector. This difference in residual inequality could also result from a variety of sources. First, according to the human capital theory, residual inequality is often interpreted as reflecting the return to and the dispersion of unobserved skills. Compared with the state sector, the private sector is more directly exposed to market competition, under which variation in unmeasured skills is more likely to translate into different rates of pay. Also, workers in the private sector may be more heterogeneous in terms of unobserved skills than workers in the state sector (Wu and Xie 2003), which would lead to greater inequality in the former even if returns to unobserved skills were identical between the two sectors. Second, as noted earlier, state-owned firms in China enjoy a wide array of institutional protections——such as government-granted monopoly and politically-favored bank loans——that help maintain a relatively low level of earnings dispersion among their employees. Finally, the difference in residual inequality between the two sectors could also stem from their differences in occupational and industrial structure. An assessment of these competing explanations, however, requires a large dataset that includes comprehensive measures of skills and detailed occupational characteristics. I leave this challenge for future research. This study, though, has demonstrated an important micro-macro nexus, that is, given that residual inequality is higher in the private sector than in the state sector, a decline in state sector employment will raise earnings inequality in the aggregate.

Earnings inequality in urban China has been on a steady rise since the early 1980s (Jansen and Wu 2012). Unfortunately, the time span of my data does not allow me to analyze the trends prior to 1996. Nonetheless, previous research has shown that the growth in earnings inequality among urban workers up to the mid-1990s was chiefly propelled by widening regional gaps and increases in residual variation (Hauser and Xie 2005). Since then, however, the composition of the urban labor force has been significantly reshaped by college expansion, state sector downsizing, and a surge in rural-urban migration. By explicitly taking into account these policy, institutional, and demographic changes, I have shown that the growth in earnings inequality during the past one and a half decades is mainly due to composition effects and increased returns to education. In light of these results, I believe that the rise of inequality in urban China has been driven by different forces during different stages of the economic reform. Understanding such stage-dependent dynamics of earnings inequality greatly enriches our knowledge about the multifaceted processes of economic transformation in post-socialist China.

References

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature* 40(1):7–72.

Allison, Paul D. 1978. "Measures of Inequality." *American Sociological Review* 43(6):865–80.

Bandelj, Nina, and Matthew C. Mahutga. 2010. "How Socio-Economic Change Shapes Income Inequality in Post-Socialist Europe." *Social Forces* 88(5):2133–61.

Bian, Yanjie, and John R. Logan. 1996. "Market Transition and the Persistence of Power: The Changing Stratification System in Urban China." *American Sociological Review* 61(5):739–58.

Cao, Yang, and Victor G. Nee. 2000. "Comment: Controversies and Evidence in the Market Transition Debate." *The American Journal of Sociology* 105(4):1175–89.

Cao, Yuanzheng, Yingyi Qian, and Barry R. Weingast. 2003. "From Federalism, Chinese Style to Privatization, Chinese Style." *Economics of Transition* 7(1):103–31.

Chan, Kam Wing, and Will Buckingham. 2008. "Is China Abolishing the Hukou System?" *The China Quarterly* 195(6):582–606.

Démurger, Sylvie et al. 2002. "Geography, Economic Policy, and Regional Development in China*." *Asian Economic Papers* 1(1):146–97.

Gan, Li et al. 2014. *Data You Need to Know about China*. Springer-Verlag Berlin Heidelberg.

Golley, Jane. 2002. "Regional Patterns of Industrial Development during China's Economic Transition." *Economics of Transition* 10(3):761–801.

Hauser, Seth M., and Yu Xie. 2005. "Temporal and Regional Variation in Earnings Inequality: Urban China in Transition between 1988 and 1995." *Social Science Research* 34(1):44–79.

Hvistendahl, Mara. 2013. "The Numbers Game." *Science* 340(6136):1037–39.

Jacobs, David. 1985. "Unequal Organizations or Unequal Attainments? An Empirical Comparison of Sectoral and Individualistic Explanations for Aggregate Inequality." *American Sociological Review* 166–80.

Jansen, Wim, and Xiaogang Wu. 2012. "Income Inequality in Urban China, 1978-2005." *Chinese Sociological Review* 45(1):3–27.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101:410–42.

Kanbur, Ravi, and Xiaobo Zhang. 2005. "Fifty Years of Regional Inequality in China: A Journey Through Central Planning, Reform, and Openness." *Review of Development Economics* 9(1):87–106.

Kim, ChangHwan, and Arthur Sakamoto. 2008. "The Rise of Intra-Occupational Wage Inequality in the United States, 1983 to 2002." *American Sociological Review* 73(1):129–57.

Knight, John B., and Richard H. Sabot. 1983. "Educational Expansion and the Kuznets Effect." *The American Economic Review* 73(5):1132–36.

Knight, John, and Lina Song. 1999. *The Rural-Urban Divide: Economic Disparities and Interactions in China*. Oxford University Press.

Lai, Qing. 2014. "Chinese Adulthood Higher Education: Life Course Dynamics under State Socialism." *Chinese Sociological Review*.

Lam, David, and Deborah Levison. 1992. "Age, Experience, and Schooling: Decomposing Earnings Inequality in the United States and Brazil." *Sociological Inquiry* 62(2):220–45.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *The American Economic Review* 96(3):461–98.

Li, Lulu, and Weidong Wang. 2012. "Chinese General Social Survey, 2003-2011." in *International Symposium on Longitudinal Survey Methods for Human Capital*.

Li, Shi, Hiroshi Sato, and Terry Sicular. 2013. *Rising Inequality in China*. Cambridge University Press.

Massey, Douglas S., and Deborah S. Hirst. 1998. "From Escalator to Hourglass: Changes in the U.S. Occupational Wage Structure 1949–1989." *Social Science Research* 27(1):51–71.

McCall, Leslie, and Christine Percheski. 2010. "Income Inequality: New Trends and Research Directions." *Annual Review of Sociology* 36:329–47.

Meng, Xin, Kailing Shen, and Sen Xue. 2013. "Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988--2009." *Journal of Comparative Economics* 41(1):227–44.

Meng, Xin, and Junsen Zhang. 2001. "The Two-Tier Labor Market in Urban China." *Journal of Comparative Economics* 29(3):485–504.

Mouw, T., and a. L. Kalleberg. 2010. "Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s." *American Sociological Review* 75(3):402–31.

National Bureau of Statistics of China. 2011. *China Statistical Yearbook 2011.* Beijing, China: China Statistics Press.

Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54(5):663–81.

Nee, Victor. 1991. "Social Inequalities in Reforming State Socialism: Between Redistribution and Markets in China." *American Sociological Review* 56(3):267–82.

Nee, Victor. 1996. "The Emergence of a Market Society: Changing Mechanisms of Stratification in China." *American Journal of Sociology* 101(4):908–49.

Nielsen, Francois, and Arthur S. Alderson. 1997. "The Kuznets Curve and the Great U-Turn: Income Inequality in US Counties, 1970 to 1990." *American Sociological Review* 12–33.

Peng, Yusheng. 1992. "Wage Determination in Rural and Urban China: A Comparison of Public and Private Industrial Sectors." *American Sociological Review* 57:198–213.

Sicular, Terry, Ximing Yue, Bjorn Gustafsson, and Shi Li. 2007. "The Urban–rural Income Gap and Inequality in China." *Review of Income and Wealth* 53(1):93–126.

Treiman, Donald J. 2012. "The 'Difference between Heaven and Earth': Urban–rural Disparities in Well-Being in China." *Research in Social Stratification and Mobility* 30(1):33–47.

Treiman, Donald J., and Andrew G. Walder. 1998. *Life Histories and Social Change in Contemporary China Provisional Codebook*.

UNU-WIDER. 2008. *World Income Inequality Database, Version 2.0c*.

Walder, Andrew G. 1990. "Economic Reform and Income Distribution in Tianjin, 1976–1986." Pp. 135–56 in *Chinese Society on the Eve of Tiananmen*, edited by D. Davis and E.F. Vogel. Cambridge, MA: Harvard University Press.

Wan, Guanghua. 2007. "Understanding Regional Poverty and Inequality Trends in China: Methodological Issues and Empirical Findings." *Review of Income and Wealth* 53(1):25–34.

Western, Bruce, and Deirdre Bloome. 2009. "Variance Function Regressions for Studying Inequality." *Sociological Methodology* 39(1):293–326.

Whyte, Martin K., and William L. Parish. 1985. *Urban Life in Contemporary China*. University of Chicago Press.

World Bank. 2011. *A Break with History: Fifteen Years of Inequality Reduction in Latin America*. Washington DC: USA.

Wu, Xiaogang. 2002. "Work Units and Income Inequality: The Effect of Market Transition in Urban China." *Social Forces* 80(3):1069–99.

Wu, Xiaogang, and Donald J. Treiman. 2004. "The Household Registration System and Social Stratification in China: 1955–1996." *Demography* 41(2):363–84.

Wu, Xiaogang, and Donald J. Treiman. 2007. "Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility." *American Journal of Sociology* 113(2):415–45.

Wu, Xiaogang, and Yu Xie. 2003. “Does the Market Pay off? Earnings Returns to Education in Urban China.” *American Sociological Review* 68(3):425.

Xie, Yu, and Emily Hannum. 1996. “Regional Variation in Earnings Inequality in Reform-Era Urban China.” *American Journal of Sociology* 101(4):950–92.

Xie, Yu, and Xiaogang Wu. 2008. “Danwei Profitability and Earnings Inequality in Urban China.” *The China quarterly* 195(2003):558–81.

Yang, Tao Dennis, and Hao Zhou. 1999. “Rural-Urban Disparity and Sectoral Labour Allocation in China.” *The Journal of Development Studies* 35(3):105–33.

Zhang, Zhuoni, and Xiaogang Wu. 2010. “Rural Labor Migration and Regional Disparities in Earnings Inequality in China: Evidence from the 2005 National Population Sample Survey.” *Population and Development (in Chinese)* 16(1):11–18.

Zhao, Wei, and Xueguang Zhou. 2002. “Institutional Transformation and Returns to Education in Urban China: An Empirical Assessment.” *Research in Social Stratification and Mobility* 19:339–75.

Zhou, Xueguang. 2000. “Economic Transformation and Income Inequality in Urban China: Evidence from Panel Data.” *American Journal of Sociology* 105(4):1135–74.

Notes

[1] "Counterfactual analysis" used in this study is a demographic technique and should not be construed as causal inference.

[2] In this step, the sample size dropped more substantially for CGSS 2010 than for LHSC 1996. This is mainly due to their differences in fieldwork implementation rather than a substantial decline in labor force participation. According to data from the World Bank, the labor force participation rate in China dropped by only 4 percentage points during this period, from 75% in 1996 to 71% in 2010.

[3] In LHSC 1996, profits from private businesses were measured at the family-level. Hence, I divided them by the number of working family members before treating them as a part of personal earnings.

[4] Collective institutions and enterprises typically do not receive financial support from the central and local governments. Compared with state-owned organizations, they are less regulated by the state and closer to market forces. Therefore, they are classified into the private sector in this study.

[5] As both estimated coefficients are asymptotically normal and independent, it is easy to show that the z-score for their difference, i.e., $\frac{\hat{\beta}_2-\hat{\beta}_1}{\sqrt{\widehat{se}^2(\hat{\beta}_2)+\widehat{se}^2(\hat{\beta}_1)}}$, is highly significant.

[6] Because there is no strong reason to assume differences in residual inequality across other social dimensions, sector of employment is used as the only predictor in the variance model.

[7] Since the college expansion primarily benefitted the younger cohorts, age and education are closely correlated in the 2010 data. Therefore, the reweighting of the educational distribution inevitably altered the age structure, which may have biased the results. To alleviate this concern, I conducted auxiliary analyses by adjusting the *conditional distribution of education given age* (i.e., $\pi_{edu|age}$) such that the educational distribution resembles that in 1996 but the age distribution remains at the 2010 level. The results are substantively identical to those reported in Table 6.

Figure 1: Earnings Inequality among Working Population in Urban China, 1996-2010

Note: Data are from the 1996 survey of "Life History and Social Changes in Contemporary China" (LHSC) and five waves of the Chinese General Social Survey (CGSS) from 2003 to 2010. Assuming the log-normality of earnings distribution, the Gini coefficients were calculated using the parametric formula $G = 2\Phi([V/2]^{0.5}) - 1$, where V is the variance of log earnings (see Allison [1978], 874).

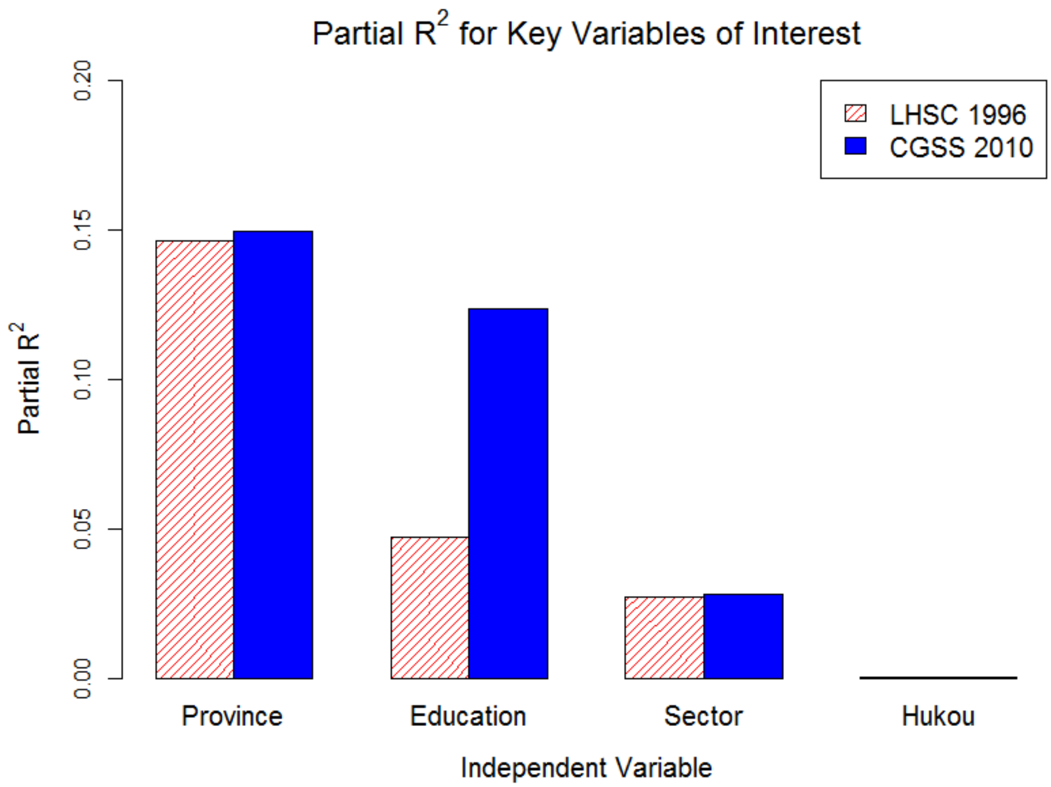

Figure 2: Partial R^2s for Province, Education, Sector, and *Hukou* Status in 1996 and 2010

Note: Besides these four key independent variables, all regression models also include sex, age, age squared, and party membership as covariates. For a variable K, partial $R^2 = \frac{R^2 - R^2_{-K}}{1 - R^2_{-K}}$, where R^2 is for the model that includes all independent variables, and R^2_{-K} is for the model that includes all independent variables except K.

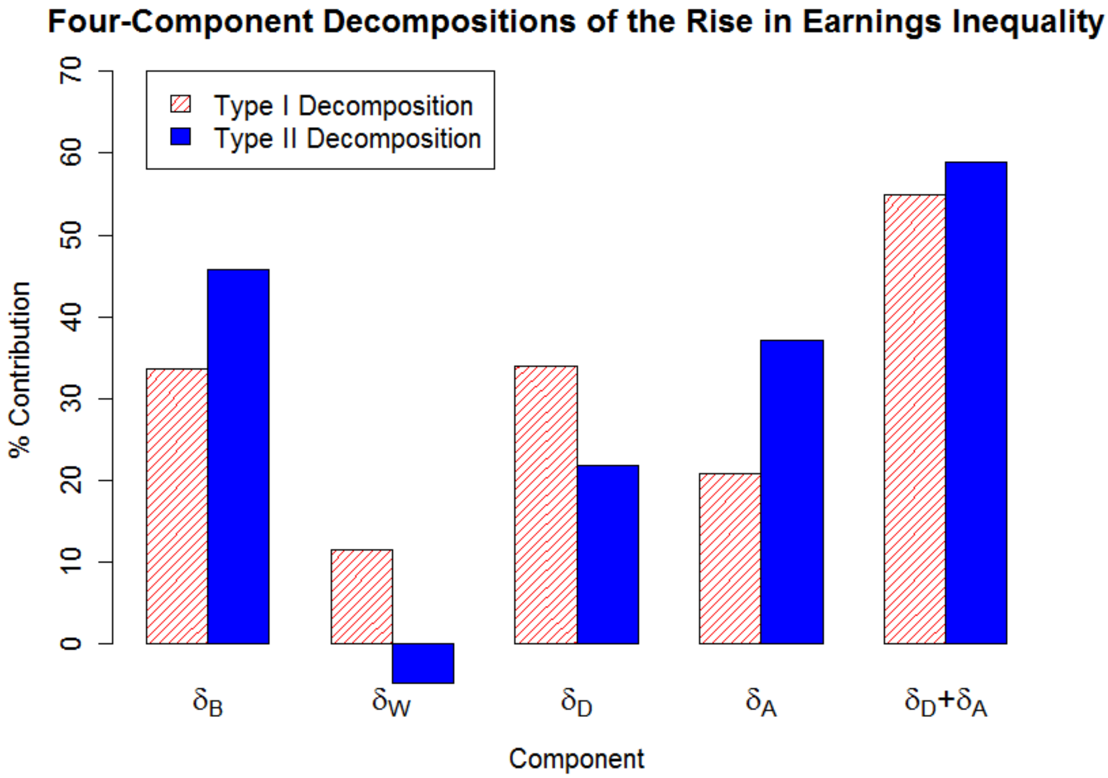

Figure 3: Decompositions of the Rise in Earnings Inequality in Urban China, 1996-2010

Note: δ_B=changes in between-group earnings gaps, δ_W=changes in within-group earnings variation, δ_D=distribution effect (δ_D), δ_A=allocation effect (δ_A), $\delta_D + \delta_A$=total composition effect.

Table 1: Four-component Decomposition of the Change in Inequality

	Changes in Between-group/Explained Inequality ($\boldsymbol{\delta_B + \delta_D}$)	Changes in Within-group/Residual Inequality ($\boldsymbol{\delta_W + \delta_A}$)
Non-compositional Changes ($\boldsymbol{\delta_B + \delta_W}$)	Changes in Between-group Earnings Gaps (δ_B)	Changes in Within-group Earnings Variation (δ_W)
Compositional Changes ($\boldsymbol{\delta_D + \delta_A}$)	Distribution Effect (δ_D)	Allocation Effect (δ_A)

Table 2: Evaluation of Competing Explanations

Competing Explanations	Mechanisms	Parameters to be Fixed at the 1996 Level
Widening Regional Disparities	Changes in Between-group Gaps	β_{region}
Increasing Returns to Education	Changes in Between-group Gaps	β_{edu}
Growing Residual Inequality	Changes in Within-group Variation	λ
Expansion of Tertiary Education	Distribution Effect	π_{edu}
Shrinkage of State Sector Employment	Allocation Effect	π_{sector}
Rural-to-urban Migration	Distribution Effect	π_{hukou}

Note: π_{edu}, π_{sector} and π_{hukou} denote the population distribution respectively by educational attainment, by sector of employment, and by *hukou* status. In this article, *hukou* status is used to distinguish between permanent urban residents and rural-urban migrants in urban China.

Table 3: Descriptive Statistics of Population Share, Mean, and Variance of Log Earnings

		Population Share		Mean of Log Earnings		Variance of Log Earnings	
		1996	2010	1996	2010	1996	2010
Sex	Male	0.59	0.57	8.94	9.97	0.56	0.85
	Female	0.41	0.43	8.65	9.63	0.51	0.80
Age	20-29	0.27	0.19	8.82	9.92	0.64	0.73
	30-39	0.31	0.31	8.79	9.97	0.42	0.81
	40-49	0.28	0.34	8.84	9.77	0.53	0.79
	50-59	0.12	0.13	8.94	9.61	0.53	0.99
	60-69	0.03	0.03	8.49	9.10	1.33	1.19
Party Membership	Not Party-member	0.82	0.81	8.77	9.73	0.60	0.84
	Party-member	0.18	0.19	9.06	10.21	0.30	0.74
Hukou Status	Urban	0.88	0.73	8.83	9.92	0.53	0.87
	Rural	0.12	0.27	8.72	9.56	0.73	0.73
Educational Attainment	No Schooling	0.03	0.02	8.13	8.89	0.56	0.75
	Elementary School	0.14	0.14	8.64	9.28	0.82	0.77
	Junior High School	0.39	0.25	8.80	9.51	0.63	0.72
	Senior High School or Vocational High School	0.30	0.27	8.87	9.82	0.36	0.66
	Vocational College	0.08	0.18	9.02	10.16	0.32	0.51
	Four-year College or Above	0.05	0.14	9.25	10.61	0.21	0.64
Sector of Employment	State Sector	0.59	0.27	8.91	10.08	0.24	0.53
	Private Sector	0.23	0.51	8.81	9.71	0.59	0.80
	Self-employment	0.18	0.23	8.52	9.77	1.43	1.26

Note: Samples sizes are 2019 and 3040 for LHSC 1996 and CGSS 2010, respectively. All numbers in this table were adjusted using sampling weights.

Table 4: Regression Results for Mean and Variance Functions in 1996 and 2010

	Explanatory Variables	Mean Regression 1996	Mean Regression 2010	Variance Regression 1996	Variance Regression 2010
	Intercept	8.690***	9.135***	-1.657***	-1.125***
		(0.158)	(0.193)	(0.064)	(0.067)
	Female	-0.222***	-0.307***		
		(0.025)	(0.026)		
	Age	0.027***	0.065***		
		(0.008)	(0.009)		
	$Age^2/100$	-0.025**	-0.089***		
		(0.009)	(0.011)		
	Party Membership	0.075*	0.155***		
		(0.031)	(0.034)		
	Rural *Hukou*	0.015	0.000		
		(0.050)	(0.034)		
Educational Attainment	No Schooling	-0.600***	-0.743***		
		(0.086)	(0.096)		
	Elementary School	-0.152***	-0.486***		
		(0.045)	(0.047)		
	Junior High School	-0.068*	-0.252***		
		(0.029)	(0.037)		
	Senior High School or Vocational High School (Reference Group)				
	Vocational College	0.079†	0.315***		
		(0.043)	(0.038)		
	Four-year College or Above	0.264***	0.608***		
		(0.051)	(0.042)		
Sector of Employment	State Sector (Reference Group)				
	Private Sector	-0.112**	-0.127***	0.794***	0.326***
		(0.036)	(0.029)	(0.121)	(0.082)
	Self-employment	-0.358***	0.183***	1.843***	1.043***
		(0.061)	(0.046)	(0.134)	(0.098)
	Model R^2	0.240	0.415		

Note: †p<.1, *p<.05, **p<.01, ***p<.001 (two-tailed tests). Numbers in parentheses are standard errors. The mean models also control for province dummies, for which the coefficient estimates are not reported here. The mean and variance models were jointly fitted via maximum likelihood estimation (Western and Bloome 2009).

Table 5: Adjusted Variances for Changes in Between-group Gaps and Within-group Variation

	2010	Change from 1996 to 2010	Percentage of Change Explained
Fitted Variance	0.839 (0.028)	0.304 (0.044)	
Fixing Changes in			
Regional Disparities ($\boldsymbol{\beta}_{\text{region}}$)	0.839 (0.034)	0.305 (0.041)	-0.2 (6.7)
Returns to Education ($\boldsymbol{\beta}_{\text{edu}}$)	0.701 (0.027)	0.167 (0.042)	**45.2** **(7.5)**
All Between-group Gaps ($\boldsymbol{\beta}$)	0.699 (0.030)	0.165 (0.038)	**45.8** **(7.9)**
All Within-group Variation ($\boldsymbol{\lambda}$)	0.853 (0.046)	0.319 (0.032)	-4.7 (16.1)

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.

Table 6: Adjusted Variances for Changes in Population Composition

	2010	Change from 1996 to 2010	Percentage of Change Explained		
			Distribution Effect	Allocation Effect	Total
Fitted Variance	0.839 (0.028)	0.304 (0.044)			
Fixing Compositional Changes in					
Hukou Status ($\boldsymbol{\pi}_{\text{hukou}}$)	0.826 (0.030)	0.292 (0.044)	-1.5 (1.2)	5.6 (1.4)	4.1 (2.4)
Education ($\boldsymbol{\pi}_{\text{edu}}$)	0.802 (0.028)	0.268 (0.044)	**21.9** **(1.6)**	-9.8 (1.5)	12.1 (2.6)
Sector ($\boldsymbol{\pi}_{\text{sector}}$)	0.780 (0.033)	0.246 (0.046)	-1.6 (2.9)	**20.8** **(2.7)**	19.2 (5.0)
Education+ Sector ($\boldsymbol{\pi}_{\text{edu,sector}}$)	0.711 (0.035)	0.177 (0.048)	21.0 (4.5)	20.8 (4.0)	**41.9** **(7.0)**
All Explanatory Variables (X, Z)	0.672 (0.026)	0.137 (0.037)	34.0 (5.9)	20.8 (4.5)	**54.9** **(7.6)**

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.

SUPPLEMENTARY DATA

Increasing Returns to Education, Changing Labor Force Structure, and the Rise of Earnings Inequality in Urban China, 1996-2010

Compositional Changes in Urban China, 1996-2010

In the main text, we have discussed three large-scale compositional changes in the urban labor force since the mid-1990s: a rapid increase in the share of college-educated workers, a tremendous decline in state-sector employment, and an upsurge in the number of rural migrant workers. In particular, the increase in college-educated workers is mainly a result of the college expansion policy implemented in the year of 1999. Figure S1 depicts these compositional shifts using external data published by National Bureau of Statistics of China (henceforth NBS) and the World Bank. First, the solid line shows the share of people who had finished a three-year college (*dazhuan*) or above among the urban population at ages 6 and above.[1] We can see that the share of college-educated people has increased sharply since the year of 2003, the time when the first "expanded" cohort of four-year college students graduated. Second, the dashed line exhibits the declining share of state sector workers in urban China since the mid-1990s, when massive SOE downsizing began to take place. Finally, the dot-dash line shows the steep rise in the number of rural migrant workers. We can see that the size of rural migrant workers more than tripled during this period, from below 50 million in 1997 to above 150 million in 2010.

[Figure S1 here]

[1] It would be more relevant to show the trends in educational distribution among the urban labor force (rather than among the total urban population at ages 6 and above). Unfortunately, such data are not available from official publications.

Sensitivity Analyses

As mentioned in the main text, the two data sets used in this study are highly comparable in sampling design. However, a major difference between LHSC 1996 and CGSS 2010 is the way households were selected within neighborhoods/villages: while CGSS 2010 employed the street-mapping technology to select households, LHSC 1996 sampled households through the combination of the list of permanent residents (the hukou register) and the list of temporary migrants that were both obtained from the neighborhood/village committee. Hence, it is reasonable to suspect that this difference in fieldwork implementation may curtail the comparability of my samples, thus weakening the credibility of the results. For example, if rural migrant workers were disproportionately unregistered in the official lists, LHSC 1996 would be less likely to catch the migration population than CGSS 2010. In this case, the observed increase in the share of rural migrant workers may have been exaggerated. Moreover, inconsistencies in various aspects of fieldwork implementation, such as what time to conduct the interview, could also lead to differences in sample representativeness. For instance, to the extent that younger people are relatively difficult to reach at home, LHSC 1996 and CGSS 2010 may also differ in their capacity to represent the age distribution of the population.

To check the representativeness of my data, I now compare my analytical samples with a combination of different external sources in two aspects: (a) the percentage of rural migrant workers in the urban labor force, (b) the age distribution of the urban labor force between ages 20 and 69. In Table S1, the first row shows the total numbers of rural migrant workers in 1996 and 2010. Specifically, the number in 2010 is from an NBS report published in 2012. Because the exact number in 1996 is not directly available, it is approximated as the average between that in 1995 and that in 1997, which have been reported respectively by Li, Sato, and Sicular (2013) and World Bank (2009). The second row presents the sizes of regularly employed urban population in these two years, which are from China Population and Employment Statistics Yearbook 2011. Denoting these two quantities as A and B, I calculated the share of rural migrant workers in the urban labor force as $\frac{A}{A+B}$. The results are reported in the third row, together with the proportions of rural *hukou* in my analytical samples of LHSC 1996 and CGSS 2010. We can see that for both years, the sample percentages, i.e., 12.1% and 27.0%, are slightly lower than the population percentages calculated from the external sources, implying an underrepresentation of rural migrant workers in both surveys. The bottom panel compares the age distributions in my data with those published by the NBS. Again, we find slight discrepancies for both years. Specifically, in LHSC 1996, there seems to be an underrepresentation of the age group 30-39 and an overrepresentation of the age group 50-59. In CGSS 2010, the youngest group (i.e., 20-29) is apparently underrepresented, which is accompanied by an overrepresentation of the age group 40-49.

[Table S1 here]

Since different individual attributes and the changes thereof are inextricably linked together, it would be difficult to speculate how and to what extent my main conclusions might be affected by the sampling biases discussed above. To check the robustness of my results, I conducted sensitivity analyses by "calibrating" my data according to the external distributions in Table S1 and replicating the main analyses. Specifically, to assess the influence of the underrepresentation of rural migrant workers, I re-weighted my analytical samples such that the proportions of rural migrant workers are identical to those calculated from the external sources, i.e., 14.7% in 1996 and 30.7% in 2010. Similarly, to assess the effects of sampling biases in age distribution, I re-weighted the original samples such that the proportions of different age groups are the same as the official figures. The results of these two analyses are presented respectively in Tables S2-S3 and in Tables S4-S5, which correspond to Tables 5-6 in the main text. We can clearly see that my substantive findings are highly robust to these weighting adjustments. First, both Table S2 and Table S4 show that more than 40% of the growth in inequality is due to changes in between-group gaps, which are almost entirely driven by increases in returns to education. Second, both sets of results suggest that over half of the total growth in inequality is attributable to composition effects, which stem mainly from changes in educational and sectoral distributions (see Tables S3 and S5).

[Table S2-S5 here]

A caveat is in order. The rationale for the reweighting method used above is that those successfully interviewed individuals can well represent the corresponding subpopulations. This is true only if there is no unobserved selection into the sample. This assumption can be violated in reality. For example, among rural migrant workers, those with lower SES may be less likely to be interviewed because of low literacy or communication skills. In this case, the sampling bias would not be corrected by the reweighting method. However, unless LHSC 1996 and CGSS 2010 differed systematically in patterns of unobserved selectivity, my analytical results would not be contaminated in an obvious fashion.

References

Li, Shi, Hiroshi Sato, and Terry Sicular. 2013. *Rising Inequality in China*. Cambridge University Press.

National Bureau of Statistics of China. 2012. *National Monitoring Survey Report on Migrant Workers*. Retrieved (http://www.stats.gov.cn/tjsj/zxfb/201305/t20130527_12978.html).

National Bureau of Statistics of China. 1997-2011. *China Labour Statistics Yearbook*. Beijing, China: China Statistics Press.

National Bureau of Statistics of China. 1997-2011. *China Population and Employment Statistics Yearbook*. Beijing, China: China Statistics Press.

World Bank. 2009. *From Poor Areas to Poor People: China's Evolving Poverty Reduction Agenda*. Retrieved (http://siteresources.worldbank.org/CHINAEXTN/Resources/318949-1239096143906/China_PA_Report_March_2009_eng.pdf).

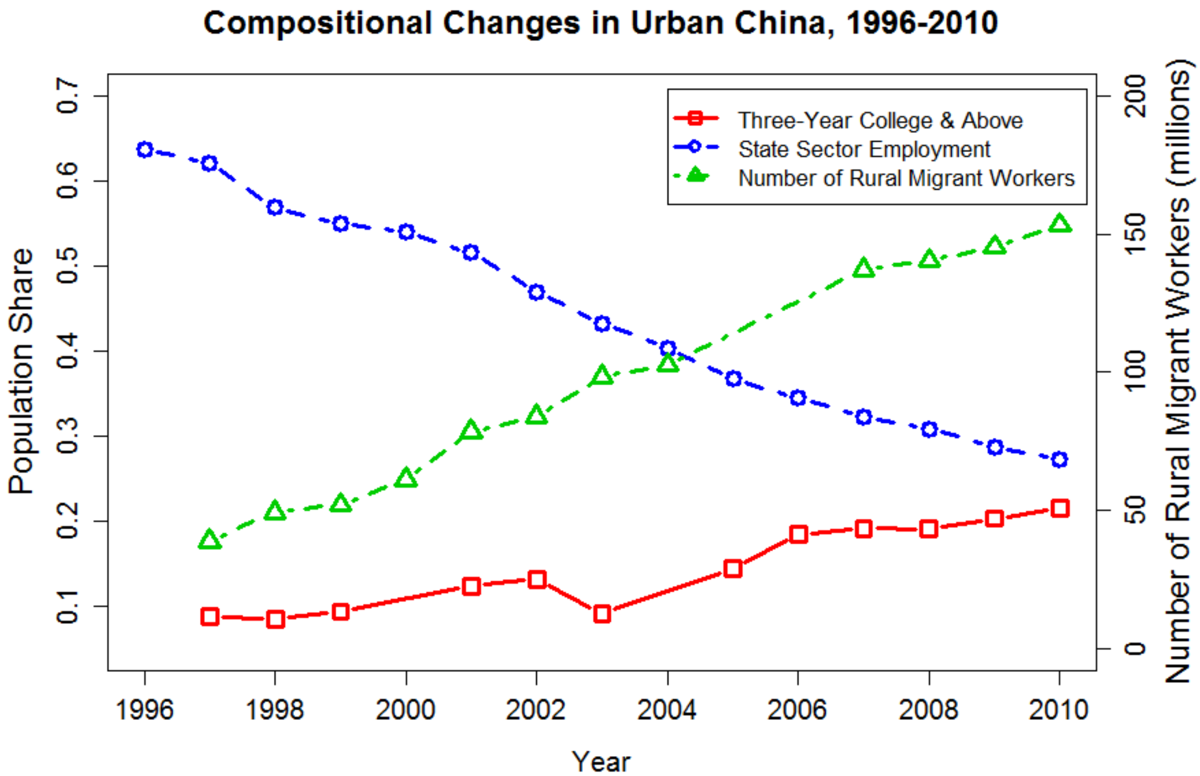

Figure S1: Compositional Changes in Urban China, 1996-2010

Note: The solid line shows the increasing share of college-educated people among the urban population at ages 6 and above (source: China Population and Employment Statistics Yearbook); the dashed line shows the declining share of workers in the state sector in urban China (source: China Labour Statistical Yearbook); the dot-dash line shows the increasing numbers of rural migrant workers in urban China (source: World Bank [2009] for years 1997-2007 and National Bureau of Statistics of China for years 2008-2010).

Table S1: Comparison of External Sources and the Author's Data on Proportion of Rural Migrant Workers and Age Distribution in the Urban Labor Force.

		1996		2010	
		External Data	LHSC 1996	External Data	CGSS 2010
Number of Rural Migrant Workers in Urban Areas (millions)		34.45		153.35	
Size of Regularly Employed Population in Urban Areas (millions)		199.22		346.87	
Proportion of Rural Migrant Workers in the Urban Labor Force (%)		14.7	12.1	30.7	27.0
Proportions of Different Age Groups among the Urban Employed Population at Ages 20-69 (%)	20-29	27.7	26.9	25.7	18.9
	30-39	33.6	30.6	30.1	31.0
	40-49	27.5	27.6	28.6	34.2
	50-59	9.3	12.0	12.3	12.9
	60-69	1.9	3.0	3.3	3.0

Note: The number of rural migrant workers in 1996 (34.45 million) is approximated as the average of that in 1995 (30 million, Li, Sato, and Sicular [2013]) and that in 1997 (38.9 million, Work Bank [2009]). The number of rural migrant workers in 2010 is from the official website of National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/zxfb/201305/t20130527_12978.html). The sizes of regularly employed population in urban China are from China Population and Employment Statistics Yearbook 2011. Denoting these two quantities as A and B, the proportion of rural migrant workers in the urban labor force was calculated as A/(A+B). Data on age distributions are from China Labour Statistics Yearbook 1999 and 2011, which report age distributions among the urban employed population in 1998 and 2010, respectively. The age distribution in 1998 is used to approximate that in 1996 because the corresponding data are unavailable for 1996 and 1997.

Table S2: Adjusted Variances for Changes in Between-group Gaps and Within-group Variation When Data are Reweighted According to External Sources on *Hukou* Composition

	2010	Change from 1996 to 2010	Percentage of Change Explained
Fitted Variance	0.836 (0.029)	0.295 (0.041)	
Fixing Changes in			
Regional Disparities ($\boldsymbol{\beta}_{\mathbf{region}}$)	0.840 (0.032)	0.299 (0.039)	-1.3 (6.0)
Returns to Education ($\boldsymbol{\beta}_{\mathbf{edu}}$)	0.702 (0.030)	0.161 (0.040)	**45.5** **(7.6)**
All Between-group Gaps ($\boldsymbol{\beta}$)	0.702 (0.032)	0.162 (0.037)	**45.3** **(8.4)**
All Within-group Variation ($\boldsymbol{\lambda}$)	0.847 (0.047)	0.307 (0.034)	-3.9 (15.7)

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.

Table S3: Adjusted Variances for Changes in Population Composition When Data are Reweighted According to External Sources on *Hukou* Composition.

	2010	Change from 1996 to 2010	Percentage of Change Explained		
			Distribution Effect	Allocation Effect	Total
Fitted Variance	0.836 (0.029)	0.295 (0.041)			
Fixing Compositional Changes in					
Hukou Status ($\boldsymbol{\pi}_{\mathbf{hukou}}$)	0.824 (0.032)	0.284 (0.042)	-2.1 (1.6)	6.1 (2.0)	3.9 (3.4)
Education ($\boldsymbol{\pi}_{\mathbf{edu}}$)	0.798 (0.029)	0.258 (0.042)	**22.3** **(1.7)**	-9.5 (1.8)	12.7 (3.1)
Sector ($\boldsymbol{\pi}_{\mathbf{sector}}$)	0.784 (0.034)	0.243 (0.044)	-3.0 (2.8)	**20.7** **(3.1)**	17.7 (5.1)
Education+ Sector ($\boldsymbol{\pi}_{\mathbf{edu,sector}}$)	0.713 (0.035)	0.172 (0.047)	20.9 (4.0)	20.7 (4.4)	**41.6** **(6.9)**
All Explanatory Variables ($\boldsymbol{X}$)	0.673 (0.026)	0.133 (0.035)	34.4 (6.1)	20.7 (4.7)	**55.1** **(7.6)**

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.

Table S4: Adjusted Variances for Changes in Between-group Gaps and Within-group Variation When Data are Reweighted According to External Sources on Age Distribution.

	2010	Change from 1996 to 2010	Percentage of Change Explained
Fitted Variance	0.837 (0.033)	0.314 (0.049)	
Fixing Changes in			
Regional Disparities ($\boldsymbol{\beta}_{\mathbf{region}}$)	0.837 (0.036)	0.314 (0.049)	-0.0 (5.9)
Returns to Education ($\boldsymbol{\beta}_{\mathbf{edu}}$)	0.710 (0.030)	0.187 (0.047)	**40.3** **(7.6)**
All Between-group Gaps ($\boldsymbol{\beta}$)	0.700 (0.032)	0.177 (0.044)	**43.6** **(8.3)**
All Within-group Variation ($\boldsymbol{\lambda}$)	0.831 (0.044)	0.308 (0.032)	1.7 (16.1)

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.

Table S5: Adjusted Variances for Changes in Population Composition When Data are Reweighted According to External Sources on Age Distribution.

	2010	Change from 1996 to 2010	Percentage of Change Explained		
			Distribution Effect	Allocation Effect	Total
Fitted Variance	0.837 (0.033)	0.314 (0.049)			
Fixing Compositional Changes in					
Hukou Status ($\boldsymbol{\pi}_{\mathbf{hukou}}$)	0.824 (0.034)	0.301 (0.050)	-1.4 (1.1)	5.4 (1.7)	4.0 (2.5)
Education ($\boldsymbol{\pi}_{\mathbf{edu}}$)	0.802 (0.033)	0.279 (0.050)	**21.0** **(1.4)**	-9.9 (1.6)	11.1 (2.5)
Sector ($\boldsymbol{\pi}_{\mathbf{sector}}$)	0.782 (0.037)	0.259 (0.052)	-1.3 (2.8)	**18.6** **(3.0)**	17.4 (5.0)
Education+ Sector ($\boldsymbol{\pi}_{\mathbf{edu,sector}}$)	0.715 (0.043)	0.193 (0.055)	20.0 (4.5)	18.6 (5.1)	**38.6** **(8.1)**
All Explanatory Variables ($\boldsymbol{X}$)	0.670 (0.028)	0.148 (0.043)	34.4 (6.5)	18.6 (5.0)	**53.0** **(8.8)**

Note: Numbers in parentheses are bootstrap standard errors (250 replications). Bolded numbers identify the main driving forces of the rise in inequality.