Inequality of Opportunity in Earnings in Rural China

X. Shi;

Australian National University, Australian Centre on China in the World , Australia

Corresponding author email: xinjie.shi@anu.edu.au

Abstract:

This paper seeks to quantify the role of inequality of opportunity in individual earnings that is associated with family background, gender, ethnic minority status, region of birth and birth cohorts in rural China. Using the China Labour-force Dynamics Survey (CLDS) for 2014, we find that the share of inequality of opportunity in individual earnings in rural China for the full sample is 20.4 percent. A Shapley-value decomposition approach reveals the contribution of each of the circumstances. This result varies across birth cohorts: the youngest cohort 1981-1990 has the lowest total inequality in earnings, but it turns out to be the one with highest circumstantial inequality as well as the partial inequality of opportunities stemming from each of the circumstances, with the only exception of gender. A closer investigation shows that three 'effort' variables—own education, off-farm employment and marital status—are pivotal in determining income inequality, but migration is not. Circumstances influence individual earnings, not only directly, but also indirectly through these three 'effort' variables.

Acknowledegment: The author is most grateful for the technical assistance (sharing related Stata and R codes) provided by Dr. Francisco H. G. Ferreira at the World Bank, Professor Markus Jäntti at Stockholm University in Sweden and Dr. Florian Chávez Juárez at the National Laboratory for Public Policy in Mexico City,and also for the insightful comments made by Dr. Jane Golley at the Australian National University.

JEL Codes: J08, Z13

#587

Inequality of Opportunity in Earnings in Rural China

Abstract

This paper seeks to quantify the role of inequality of opportunity in individual earnings that is associated with family background, gender, ethnic minority status, region of birth and birth cohorts in rural China. Using the China Labour-force Dynamics Survey (CLDS) for 2014, we find that the share of inequality of opportunity in individual earnings in rural China for the full sample is 20.4 percent. A Shapley-value decomposition approach reveals the contribution of each of the circumstances. This result varies across birth cohorts: the youngest cohort 1981-1990 has the lowest total inequality in earnings, but it turns out to be the one with highest circumstantial inequality as well as the partial inequality of opportunities stemming from each of the circumstances, with the only exception of gender. A closer investigation shows that three 'effort' variables—own education, off-farm employment and marital status—are pivotal in determining income inequality, but migration is not. Circumstances influence individual earnings, not only directly, but also indirectly through these three 'effort' variables.

1 Introduction

Since 1978, China's economic reforms have not only led to rapid economic growth—which has long been hailed as one of the biggest contributors to the decrease in the global rate of extreme poverty—but also an accompanying downside: a widening income gap. Although the economy has slowed in recent years, income inequality remains high in China. According to Xie and Zhou (2014), China's Gini coefficient nearly doubled from around 0.30 in 1980 to 0.55 in 2012. There may be no other case where a nation's income distribution has deteriorated so much and so dramatically (Naughton, 2007). A rapidly growing literature has examined the determinants of China's high income inequality, of which a substantial part is due to regional disparities (Gustafsson & Li, 2002) and the rural-urban gap, which is linked to the *hukou* system (Li & Gibson, 2013; Sicular et al., 2007; Yang, 1999). Other factors include gender(Chen et al., 2013; Matthews & Nee, 2000), political status (Morduch & Sicular, 2000; Walder, 2002), and human capital(Fleisher, Li & Zhao, 2010).

Inequality in China is clearly far more complex than just the rural-urban dichotomies. Even within rural areas, inequality between villages and within villages has increased greatly (Zhou, Han & Harrell, 2008). There is an extensive literature focusing on rural China, examining two main issues(Benjamin, Brandt & Giles, 2005): (i) the estimation of the level of and changes in inequality, and (ii) the key factors contributing to those trends. On the first issue, despite the diverse data used in the literature, there is a rough consensus that rural inequality in China has increased during the reform period (Benjamin, Brandt & Giles, 2005; Wan & Zhou, 2005). Bonnefond and Clément (2012) conclude that rural inequality is greater than urban inequality. Turning to the second issue, many empirical studies have highlighted the crucial contribution of off-farm income and remittances to the increase in rural inequality over time (Benjamin, Brandt & Giles, 2005; Bonnefond & Clément, 2012; Howell, 2017; Kung & Lee, 2001; Scharf & Rahut, 2014). For example, Clément (2016) shows that the fall in agriculture income—which is an equalizing component of income due to the relatively egalitarian distribution of land and collective ownership—and the increase in off-farm income, account for the largest part of the increase in rural inequality.

There has been no research to my knowledge, however, on whether the income distribution in rural China is fair or not – that is, whether it constitutes equal opportunity, or 'inequality of opportunity'. A rapidly growing literature in this field addresses this issue from the basic

premis that an individual's achievements should depend only upon his efforts (and choices), and not on predetermined circumstances over which he has no control.

This idea became prominent in the late 1960s and early 1970s when researchers identified different returns arising from identical level of efforts by individuals with different backgrounds (Bowles, 1972; Hanoch, 1967; Weiss, 1970). This literature initiated a new agenda where researchers focused on how family background mattered in an individual's overall economic achievements (e.g. earnings). This kind of inequality due to circumstances (inequality of opportunity) has been highlighted and developed by a number of studies. The pioneering work of Roemer (1993, 1998) formalized the concept of unequal opportunities and separated the determinants of individual's "advantage" (i.e. desirable outcomes, such as incomes or educational attainments) into circumstances which are exogenous to the individual—for instance, their gender, place of birth and the socioeconomic status of their parents— and efforts which are under the control of the individual—for example, their own educational achievement and their choosing to work in off-farm sector in the context of rural areas.

The recent popular saying *pingdie* in Chinese –"daddy is the key"—indicates that family background (and other circumstances), instead of fair competition among individuals, has become a significant part of individual success in China. This study investigates whether this is actually the case by estimating inequality of opportunity in earnings in rural China and identifying what circumstances are crucial in determining individual earnings and what efforts are useful for eliminating income inequality. Father's education and occupation are two of the key circumstances used in the empirical literature (Ferreira & Gignoux, 2011; Peragine & Ferreira, 2015; Singh, 2012). The reason is that well-educated and wealthy parents are more likely to invest in educating their children and make use of their social network to benefit their children as they seek to enter the labour market. Regional disparities have been identified as a major contributor to income inequality due to a widening divide between industrially developing areas, mostly near the coast or large cities, and areas mainly relying on agriculture, mostly inland and far from major industrial activity (Fleisher, Li & Zhao, 2010; Li & Gibson, 2013; Sicular et al., 2007; Zhou, Han & Harrell, 2008). It is inevitable that a country as large as China—comprised of 34 provincial-level administrative regions—will have large spatial differences in socioeconomic development. So, it is essential to take into account the consequence of the vast regional difference in China. We use the

region of birth to capture the influences of geographical circumstances as most of the western and central regions are economically and demographically lagging behind the eastern regions. We also consider the income gaps in rural China between ethnic groups (with Han Chinese maintaining their traditional lead), as shown in Gustafsson and Shi (2003) and between genders (with men more likely to attain higher income than women), as shown in Matthews and Nee (2000) and Hannum (2005). The last "circumstance is one's birth cohort, which is quite important in the context of rural China as people born in different periods have faced different opportunities.

The rest of the paper proceeds as follows. Section 2 reviews the empirical applications of inequality of opportunity in a number of nations. Section 3 introduces the conceptual framework and the methodology for the estimation of inequality of opportunity. Section 4 describes the data, the specification of the outcome variable, circumstances, and efforts, and the empirical results on inequality of opportunity in earnings in China. Sections 5 further examines the role of efforts other than circumstances in determining the total inequality in earnings. Sections 6 concludes.

2 Background

In contrast to the normative literature, empirical studies on 'inequality of opportunity' are still relatively rare although growing rapidly. Building on the ideas of a number of authors (Arneson, 1989; Cohen, 1989; Dworkin, 1981) who argue that not all the differences in particular outcomes or 'advantage' (e.g. incomes, or educational attainments) are unacceptable, the economic literature partitions the observed inequality in particular economic outcomes into two components. The first component (acceptable inequalities) stems from the different factors for which individuals can be held responsible. The second component (unacceptable inequalities) –defined as 'inequality of opportunity'—is attributable to different factors over which individuals have no control.

The pioneering work of Roemer (1998) further defines these two types of factors as 'efforts' (e.g. how long one studies, or whether one chooses to be off-farm employed) and 'circumstances' (e.g. gender, ethnic minority status, or family background) respectively. He further defines 'equality of opportunity' as a situation where outcomes (advantages) are

distributed independently of circumstances. This is the main idea of the ex-ante approach[1] to measuring inequality of opportunity in which there is equality of opportunity if the set of opportunities is the same for all individuals, regardless of their circumstances. Specifically, Roemer partitions the population into 'types' formed by individuals endowed with the same set of circumstances. Hence, the overall inequality in income is decomposed into between-group and within-group components, with the former one ('between-type' inequality) being used as a measure of inequality of opportunity

Building on Roemer's ideas, many papers have estimated inequality of opportunity in different country settings, developing two main approaches to measuring opportunity inequality, of which non-parametric approach is the choice of some studies. Checchi and Peragine (2010), for instance, provide a non-parametric approach to measure opportunity inequality in Italy. They show how parental education, taken as a circumstance out of individual control, significantly affect the equality of opportunities especially when considering population subgroups (by gender and by region of residence). Lefranc, Pistolesiand Trannoy (2008) use stochastic dominance rankings of distributions conditional on types as a measure of inequality of opportunity. However, the standard non-parametric approach is only optimal when there are relatively few types. As the number of types increases, the frequency of sample observations per type tends to diminish quite rapidly. In both these two studies, inequality of opportunity is associated only with differences between 3 or 5 groups. Such a restriction demands a large sample for each type, otherwise, it is likely to lead to an underestimate of inequality of opportunity.

Largely due to data limitations, many studies use the parametric approach instead. Bourguignon, Ferreiraand Menendez (2007), for example, estimate a linear model of earnings as a function of circumstances and efforts, and use it to simulate counterfactual distributions where the effect of circumstances is suppressed. Inequality of opportunity in Brazil corresponds to a component due to five observed circumstances, which is decomposed by comparing the actual earnings distribution with different counterfactuals. Singh (2012) adopts a parametric approach to estimate inequality of opportunity in India that is associated with several circumstances, including parental education, parental occupation, caste, religion,

[1] In contrast, ,an 'ex-post' measure of inequality of opportunity corresponds to inequality within 'tranches', which is defined in terms of one's relative position in the effort distributions across types (Checchi and Peragine, 2010)

and place of birth. The overall opportunity share estimates of earnings inequality vary from 18 to 26 percent for urban India, and from 16 to 21 percent for rural India. Some other associated papers include Checchi, Peragineand Serlenga (2010) with a focus on 25 European countries, and Ferreira and Gignoux (2011), focusing on a number of Latin America countries.

To the best of our knowledge, none of these cited works on inequality of opportunity are specifically about China, with the exception of only two papers. Zhang and Eriksson (2010) use data from the China Health and Nutrition Survey collected from nine provinces during the period 1989 to 2006, to find that China has a rising degree of inequality of opportunity in incomes that is largely mirrored by the increase in income inequality. Golley and Kong (2016) measure the inequality in individual educational outcomes in aggregate and for each of ten birth cohorts, utilizing the China Family Panel Studies survey for 2010. Their results are based on a parametric approach, and show that the *hukou* system is the dominant circumstance variable in determining the educational outcomes, with father's education, birth cohort, province, parents' Communist Party membership, gender, family size and ethnicity also playing important roles. However, they have neither focused on inequality of opportunity in earnings in rural China nor add 'effort' to the framework.

Note that the discussion above mainly points to overall inequality of opportunity. Bourguignon, Ferreiraand Menendez (2007) made the first attempt at (i) decomposing the effect of opportunities into a direct effect on earnings and an indirect component, which works through the "effort" variables; and (ii) estimating the contribution of each of the circumstance variables to earnings inequality using a parametric approach. However, a corrigendum to this (Bourguignon, Ferreira & Menéndez, 2013) shows that this approach is not reliable. Björklund, Jänttiand Roemer (2011)'s more recent attempt to disentangle the direct and indirect effects of circumstances on long-run income using a Shapley-value decomposition. Jusot, Tubeufand Trannoy (2013) take a different approach, adopting three alternative ways of treating the correlation between circumstances and efforts championed by Roemer, Barry and Swift[2] to compare the relative contributions of circumstances and efforts to overall health inequality using regression analysis with the natural decomposition of the variance. Their results show little difference for these three normative principles. Although

[2] The details are shown in the next section.

this approach allows us to decompose overall inequality into circumstance-induced inequality and effort-induced inequality by including a set of circumstance and effort variables in the regression, it is not directly comparable with Bourguignon, Ferreiraand Menendez (2007) and Björklund, Jänttiand Roemer (2011). We will discuss this in detail in the next section.

To examine the overall inequality of opportunity in China's rural labour earnings, the present study adopts the ex-ante framework proposed by Ferreira and Gignoux (2011) and Bourguignon, Ferreiraand Menendez (2007). The primary reason for using parametric approach is that non-parametric approach not only suffers from data insufficiency problems once the number of circumstance variables increases, but also fail to capture the partial effects of circumstances. We also adopt the idea of Björklund, Jänttiand Roemer (2011) to identify the contribution of each of the circumstances and the approach of Jusot, Tubeufand Trannoy (2013) to enhance on the contribution of some of the important effort variables in rural China.

3 Analytical Framework

3.1 Inequality of opportunity

Consider a finite population of individuals indexed by $i \in \{1,...,\mathrm{N}\}$, each of whom has attained a set of attributes $\{\mathrm{y}_i, \mathrm{C}_i, \mathrm{E}_i\}$, where y denotes a level of income, C denotes a vector of circumstances, which lie beyond the control of the individual, and E denotes a vector of effort, which can be affected by individual choice. C_i consists of J elements corresponding to each circumstance j, and each element $C_i^{\,j}$ takes on a finite number of values, $x_j, \forall i$. This helps partition the population into K types, in which individuals have identical circumstances (Roemer, 1998), given by $\Pi \in \{\mathrm{T}_1, \mathrm{T}_2, ..., \mathrm{T}_k\}$, such that $T_1 \bigcup ... \bigcup T_k = \{1,...,\mathrm{N}\}$, $T_h \bigcap T_l = \varnothing, \forall h,l$ and $C_t = C_j, \forall t, j \mid t \in T_k, j \in T_k, \forall k$. This results in the maximum possible number of types that is given by $\overline{K} = \prod_{j=1}^{J} x_j$.

Following Roemer (1998), Bourguignon, Ferreiraand Menendez (2007), and Ferreira and Gignoux (2011), the parametric approach starts with the following function.

(1) $y = f(C, E, u)$

Where y denotes individual earnings and is a function of circumstance variables C, effort variables E and other unobserved determinants u. Note that circumstance variables are economically exogenous, but that effort itself can depend on circumstances, as well as other random factors. In this case, (1) can be rewritten as:

(2) $y = f\left(C, E(\mathrm{C}, \mathrm{v}), u\right)$

Given that the distribution of earnings is independent of circumstances $F(\mathrm{Y} \mid \mathrm{C}) = \mathrm{F}(\mathrm{Y})$, Roemer's concept of equality of opportunity would attain when the following two conditions hold (Bourguignon, Ferreira & Menendez, 2007):

(i) $\partial f(\mathrm{C}, \mathrm{E}, \mathrm{u}) \big/ \partial C = 0, \forall C$, circumstances have no direct effect on advantages conditional on efforts.

(ii) $G(\mathrm{E} \mid \mathrm{C}) = G(\mathrm{E}), \forall E, \forall C$, efforts should be distributed independently from circumstances. In other words, circumstances have no causal effect on efforts.

To measure inequality of opportunity is therefore to measure the extent to which $F(\mathrm{Y} \mid \mathrm{C}) \neq \mathrm{F}(\mathrm{Y})$. As stated in section 2, the recent literature mainly contains at least two different approaches—parametric approach and non-parametric approach—to the measurement of inequality of opportunity. This paper adopts the parametric approach. An empirically suitable first approximation can be obtained by log-linearization:

(3) $\ln(\mathrm{y}) = \alpha + \beta\,\mathrm{C} + \gamma\,\mathrm{E} + \mathrm{u}$

(4) $E = \lambda C + e$

where β and γ are two vectors of coefficients, and λ is a matrix of coefficients indicating the mechanism through circumstances affect efforts. Substituting (4) into (3) generates the reduced form of the structural model:

(5) $\ln(\mathrm{y}) = \alpha + \mathrm{C}(\beta + \gamma\lambda) + \gamma e + \mathrm{u}$

which can be simply estimated by OLS as follows:

(6) $\ln(\mathrm{y}) = \alpha + \rho\,\mathrm{C} + \varepsilon$

where ρ encompasses both the direct effect of circumstances on the advantage y, and the indirect effect of circumstances through efforts, and ε denotes the random error term. Using

the estimated coefficients ρ and the actual values of circumstances, one can construct a parametric estimate of the smoothed distribution as follows[3]:

(7) $\{y_i\}$, where $y_i = \exp[\mathrm{C}\gamma]$

Thus the parametrically smoothed (direct) estimates for inequality of opportunity indices can be defined as follows[4]:

(8) $IOA_D = I(\{y_i\})$, $IOR_D = I(\{y_i\}) / I(\{y_i\})$

We employ the general entropy index GE(0), which is decomposable and commonly used for estimating inequality of opportunity (Ferreira & Gignoux, 2011).We should always keep in mind that this method can only generate lower-bound estimates of the "true inequality of opportunity" due to the omitted variable bias.

The next step is to estimate the contribution of each of the circumstances. Bourguignon, Ferreiraand Menendez (2007) made the first attempt at doing this using a parametric approach[5]. However, a corrigendum to this (Bourguignon, Ferreira & Menéndez, 2013) shows that this approach is not reliable. Thus, we closely follow the more recent idea of Björklund, Jänttiand Roemer (2011) who use a Shapley-value decomposition (Shorrocks, 1982, 2012), which decomposes inequality of opportunity into its sources by eliminating the relative importance of each circumstance one by one. Using the Shapley-value decomposition, we first need to generate the "power set" of the K circumstances and estimate the inequality of measure for all possible permutations of these circumstances. We take every element of the power set that does not include a specific circumstance, and compare inequality in that set with the set that is otherwise identical but does include the circumstance. In a second step, the average marginal effect of each circumstance variable on the measure of inequality of opportunity is computed. Although this procedure is very computation intensive, there are two benefits in contrast to other decomposition methods. First, the

[3] The parametrically standardized distribution using γ and the mean values of each circumstance would be specified as: $\{y_{i2}\}$, where $y_{i2} = [\overline{C}\gamma + \varepsilon]$

[4] Parametrically standardized estimates are obtained as: $IOA_I = I(\{y_{i2}\}) - I(\{y_{i2}\})$, $IOR_I = I(\{y_{i2}\}) - I(\{y_{i2}\}) / I(\{y_{i2}\})$

[5] $y_j^{\ J} = \exp[\overline{C}^{i=J}\gamma^{J} + \mathrm{C}^{j\neq J}\gamma^{i\neq J} + \varepsilon]$, $IOA^J = 1 - I(\{y_j^{\ J}\})/I(\{y_j\})$,

decomposition is order independent and second, the sum of all contributions is the value of overall inequality (Juarez & Soloaga, 2014).

3.2 Adding Effort to the Framework

The method stated above has become by far dominant in the literature. But the disadvantage of this method is that the effort variables are omitted. This is likely to lead to a biased estimate of inequality of opportunity if these effort variables (or other omitted circumstances) are correlated with the included circumstances. Following the idea of Jusot, Tubeufand Trannoy (2013), we set out to address this problem by including some important effort variables in the context of rural China.

This method starts with the debate between Roemer (1998) and Barry (2005)[6]. Roemer's view is the basis of the method in Section 3.1—one's effort is clearly an aspect of the environment outside his control. An equal-opportunity policy must respect the individual effort in an approach where it can be purged of any contamination coming from circumstances. For the case that 'Asian children generally work hard in school and thereby do well because parents press them to do so', he proposes that the extra efforts of the Asian student must not be rewarded because it is determined by a characteristic which lies behind his control. This idea is laid out in equation (4), where E, the efforts of the individual, is a linear function of *C*, the circumstances. It allows isolation of a residual term *e*, the relative efforts, which represent efforts purged from circumstances. We then substitute the vector of efforts *E* for the estimated relative effort *e* in equation (3). The logarithm of earnings can be written in Roemer's perspective as follows:

(9) $\ln\hat{y} = \hat{\alpha}C + \hat{\beta}e$

Note that since the effort variables in equation (4) are binary, a Probit model is adopted. However, this does not allow us to undertake a direct estimation of the relative efforts. We thus compute generalized residuals as follows (Gourieroux et al., 1987):

(10) $E(\mathrm{e}/\mathrm{E}) = \dfrac{\varphi(\lambda\mathrm{C})}{\phi(\lambda\mathrm{C})[1-\phi(\lambda\mathrm{C})]}[\mathrm{E}-\phi(\lambda\mathrm{C})]$

[6] There is another view (Swift, 2005), which suggests that circumstances only include past variables and have to be cleaned from any correlation with descendant's effort. In the following part of this study, we just adopt the other two views.

In contrast to Roemer, Barry (2005) argues that one's efforts have to be fully respected whatever the influence of circumstance on efforts will be. In the case of 'Asian students', 'the fact that their generally high levels of effort were due to familial pressure does not make their having expended high levels of effort less admirable and less deserving than it would have been absent such pressure'. As a consequence, the extra efforts of the Asian student reflect his free will and should be entirely rewarded. This allows directly regressing circumstances and effort variables on (log) earnings, which can be written as follows in Barry's context:

(11) $\ln\hat{y} = \hat{\alpha}C + \hat{\beta}\mathrm{E}$

In contrast to the method in Section 3.1, the inequality index used here is variance. Shorrocks (1982) shows that the variance is a good index for an absolute measure of inequality and the square of the coefficient of variation is applicable for a relative measure of inequality. In this case, it does not matter whether we choose an absolute or relative inequality coefficient as the same relative decomposition for both indices applies (Jusot, Tubeuf & Trannoy, 2013). The decomposition of the variance of income is given by:

(12) $\sigma^2(y_y) = \mathrm{cov}(y_C, y_y) + \mathrm{cov}(y_{\mathrm{e/E}}, y_y)$

Where y_C represents the circumstance-related inequality, and $y_{\mathrm{e/E}}$ represents the effort-related inequality that stems from Roemer's scenario of equation (9) and Barry's scenario of (11) respectively. The contribution of circumstances and efforts can be calculated as follows:

(13) $COC = \mathrm{cov}(y_C, y_y)/\sigma^2(y_y)$, $COE = \mathrm{cov}(y_{\mathrm{e/E}}, y_y)/\sigma^2(y_y)$

4 Data and Results

4.1 Data

The data is from the China Labor-force Dynamics Survey (CLDS) produced by the Center for Social Science Survey of Sun Yat-sen University in China. CLDS is a nationally representative survey of Chinese communities, families, and individuals in contemporary China. The samples of CLDS cover 29 provinces (excluding Hong Kong, Macau and Taiwan, Tibet and Hainan), with a focus on labour in the household. Using multistage cluster, stratified PPS sampling, CLDS did the trail survey in Guangdong in 2011, conducted the first formal investigation in 2012, and completed the first tracking survey in 2014. The 2014 data has been released to the public since January 2017.

We restrict the sample to individuals born in 1951-1990 (aged 24-63) in rural areas. We choose this sample in order to focus on rural individuals with the highest levels of labour market attachment. The complete CLDS 2014 sample size is 23,594 individuals. Excluding individuals with urban hukou, and those outside 1951-1990 cohort range yields a sample of 15,035 individuals. The final sample is then divided into four 10-year birth cohorts: from individuals born between 1951-1960 through to those born between 1981-1990. This allows us to shed light on not only the role of circumstances in determining the observed earnings at one point in 2014, but also how this role may vary across cohorts.

Table 1. Earnings and Circumstance, by 10-year Cohorts

Birth Cohort	Share	Earnings				
		All	1951-1960	1961-1970	1971-1980	1981-1990
	100	27838	18097	27053	33815	33227
Gender						
Female	46	19724	10938	16925	25783	26170
Male	54	34797	23250	36010	41676	39176
Father's Education						
Illiterate	45	19639	14231	20662	25925	23823
Primary School	35	32396	21741	35195	36127	30872
Junior High School	13	37591	25690	29613	45146	38120
Senior High School	7	40496	80062	40311	33063	38168
Father's Occupation						
Low-Status Job	89	25189	17256	24302	32694	28438
High-Status Job	18	39755	25571	44218	38746	42050
Ethnic Minority Status						
Non-Han Chinese	13	18697	10623	17098	25366	21166
Han Chinese	87	29150	18994	28537	35027	35183
Birth of Region						
East	40	31422	23143	29324	37740	38060
Centre	26	29700	16529	27030	43099	35393
West	34	22040	12235	24264	24070	25399

Turning to the circumstances, the survey contains abundant information. The circumstances we choose based on the above literature include (i) gender, which is a dummy with male—who account for 54% of the respondents—taking value 1 ; (ii) ethnic minority status, which is a dummy as well, with Han taking value 1 and accounting for 87% in the full sample; (iii) father's education, which is re-coded into four categories—illiterate, primary school, junior

high school and senior high school and above; (iv) father's occupation, which is re-coded into two categories. "Lower status" includes farmers, fishermen, agricultural labourers, farm and forestry workers, hunters and related workers and "higher status" mainly includes off-farm workers, such as administrative officials, corporation leaders, professionals, clerical jobs and transport and communication supervisors; (v) region of birth, which is categorized into three regions that have been widely adopted in China: East, Central and West. The East region includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan. The central region comprises the provinces of Shanxi, Jilin, Heilongjiang, Anhui. Jiangxi. Henan, Hubei, Hunan. The West region includes Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunan, Xizang, Shan'xi, Gansu, Qinghai, Ningxia, Xinjiang. The east region is taken as the reference category.

The dependent variable is individual earnings in 2014. The preliminary statistics in Table 1 shows that the average annual earnings for rural individuals are 27, 838 yuan. This varies across birth cohorts, going up from 18, 097 yuan for the 1951-1960 birth cohort to a peak of 33, 815 yuan for 1971-1980 birth cohort. Not surprisingly, individuals born in east regions have higher income than their counterparts born in west and centre regions. Individual earnings also vary across people with other circumstances: being male and Han with more educated fathers in high-status jobs is likely to earn more in the raw data.

4.2 Determinants of individual earnings

This section briefly reports on the estimation results. Table 2 presents the regression results of the earnings equation (6) estimated by OLS, for the whole sample and separately for each cohort.

All circumstance variables have the expected sign on individual earnings for the full sample (in Column 1). The gender and ethnic biases are clear: with the females and non-Han as references, being male is associated with earnings that are 0.559 log points[7] higher and being Han is associated with earnings that are 0.265 log points higher. It is also clear that more educated father with higher income is beneficial to their children's job performance, with primary school, junior high and senior high and above being associated with earnings

[7] For the log-linear form with dummy regressor, the exact percentage change in the dependent variable (i.e., semielasticity), is (exp(β)-1)*100, where β is the coefficient of the dummy regressor. The case here amounts to 74.9%, given by (exp(0.559)-1)*100.

that are 0.225, 0.287 and 0.343 log points higher respectively compared to their counterparts with no schooling, and with high-status job of the father being associated with a 0.213 log point income boost compared to those with low-status jobs. Geographical variables are highly significant and negative, showing that being in the central and east regions is associated with earnings that are 0.282 and 0.438 log points higher than being in the west. Turning to birth cohorts, the result is consistent with the preliminary statistics that the two younger cohorts are associated with higher earnings.

Table 2. Regressions using circumstance variables across birth cohorts and regions

		Birth cohort			
	All	1951-60	1961-70	1971-80	1981-90
Male	0.559***	0.604***	0.620***	0.569***	0.392***
	(0.02)	(0.05)	(0.04)	(0.04)	(0.04)
Primary school	0.225***	0.140**	0.293***	0.183***	0.236***
	(0.03)	(0.06)	(0.04)	(0.05)	(0.07)
Junior high school	0.287***	0.291**	0.420***	0.229***	0.266***
	(0.04)	(0.14)	(0.07)	(0.07)	(0.07)
Senior high school	0.343***	0.372**	0.258**	0.229**	0.425***
	(0.05)	(0.17)	(0.11)	(0.10)	(0.08)
Father-High status job	0.213***	0.179**	0.154***	0.284***	0.261***
	(0.03)	(0.09)	(0.06)	(0.06)	(0.05)
Farther- Han Chinese	0.265***	0.395***	0.270***	0.193***	0.203***
	(0.04)	(0.09)	(0.06)	(0.07)	(0.07)
East	0.438***	0.494***	0.411***	0.471***	0.376***
	(0.03)	(0.06)	(0.05)	(0.05)	(0.06)
Centre	0.282***	0.205***	0.252***	0.346***	0.364***
	(0.03)	(0.07)	(0.05)	(0.06)	(0.06)
Birth cohort 61-70	0.479***				
	(0.03)				
Birth cohort 71-80	0.736***				
	(0.03)				
Birth cohort 81-90	0.758***				
	(0.04)				
Constant	8.188***	8.065***	8.626***	8.977***	9.069***
	(0.04)	(0.09)	(0.06)	(0.07)	(0.08)
N	9096	1993	3147	2207	1749
Adjusted R-square	0.203	0.123	0.148	0.141	0.148
F	211.588	35.812	69.175	46.165	38.997

Notes: *** 1%,** 5%,* 10%;
Standard errors in parentheses.

The regressions for each birth cohort identify important variations in the magnitude and significance of key determinants, which is shown to be the case from columns 2 to 5 in Table 2. Gender is a highly significant determinant of individual earnings for all cohorts. The coefficient of gender peaks at 0.620 log points for the cohort 1961-1970, and falls to a low of

0.392 log points the cohort 1981-1990. Other key points include the consistently positive coefficients on father's occupation and education, and the significant advantage of being born in the centre and east compared to being in the west, and of being Han in contrast to being other ethnic minorities.

4.3 Inequality of opportunity

Using the coefficient estimates from the reduced-form equation (6), reported in Table 2, we simulate the counterfactual distributions, following the method presented in Section 3.1. This helps to decompose earnings inequality for the whole sample and separately for each cohort into a component due to unequal circumstances over which individuals have no control (inequality of opportunity) and a residual component due to "efforts", random term and some unobserved circumstances.

Table 3 shows the IOA and IOR in total observed earnings inequality in rural China. We begin with the income inequality index, GE(0), and the Gini coefficient, for comparative purposes. Total inequality in earnings is high in rural China, with the GE(0) and Gini coefficient being 0.560 and 0.646 respectively. With respect to birth cohorts, total income inequality for the younger cohorts are lower than that for the older cohorts.

Table 3. Inequality of opportunity in earnings in rural China

	All	Birth Cohorts			
		1951-60	1961-70	1971-80	1981-90
Panel A					
Total inequality					
Gini	0.560	0.610	0.575	0.541	0.459
GE(0)	0.646	0.757	0.670	0.595	0.422
Inequality of Opportunity					
Absolute (IOA)	0.132	0.086	0.092	0.086	0.068
Relative Share (IOR)	**0.204**	**0.114**	**0.138**	**0.144**	**0.161**
Panel B					
Shapley decomposition (% of IOR)					
Gender	27.25%	48.06%	51.09%	46.03%	28.92%
	5.6%	**5.4%**	**7.1%**	**6.6%**	**4.7%**
Father's education	15.23%	7.69%	16.37%	8.07%	15.68%
	3.1%	**0.9%**	**2.2%**	**1.2%**	**2.5%**
Father's occupation	8.33%	4.09%	4.57%	11.17%	19.17%
	1.7%	**0.5%**	**0.6%**	**1.6%**	**3.1%**
Ethnic	5.38%	11.14%	8.42%	7.01%	10.15%
	1.1%	**1.3%**	**1.2%**	**1.0%**	**1.6%**
Birth region	14.27%	29.02%	19.55%	27.72%	26.08%
	2.9%	**3.3%**	**2.7%**	**4.0%**	**4.2%**
Birth cohort	29.54%				
	6.0%				

Turning to the inequality due to unequal opportunities, the overall share for the full sample is 20.4 percent. The IORs are much lower for each of the birth cohorts. This is due to the exclusion of birth cohort as a circumstance in estimating IORs across cohorts. The key point, however, is that IORs tend to be higher for the younger cohorts: with an IOR value of 0.161 for the 1981-1990 cohort and with an IOR value of 0.114 for the 1951-1960 cohort. Along with the results above, this indicates that the younger cohorts have higher inequality of opportunities although the total income inequality is much lower for them in contrast to the older cohorts.

To understand the contribution of each of the circumstances, we adopt a Shapley-value decomposition discussed in Section 3.1 to decompose the overall inequality of opportunity into circumstance specific parts. Panel B in Table 3 presents two types of shares for each of the circumstances—the contribution of a specific circumstance to IORs in the first row and to total income inequality in the second row. This suggests that the lack of equal opportunity for individuals in rural areas with regard to their earnings stems from birth cohorts and gender that accounts for 6.0 percent and 5.6 percent of the total income inequality respectively, with further father's education (3.1%), birth region (2.9%), father's occupation (1.7%) and ethnic (1.1%), in that order.

In a further step, we get rid of birth cohort, across which we examine the contribution of other circumstances. Gender is still dominant in contributing to the total income inequality across all the cohorts, but the relative share is much lower for the younger cohorts: 4.7 percent for the cohort 1981-1990 compared to 7.1 percent for the cohort 1961-1970. The key point, however, is that all the circumstances other than gender contribute most to the total income inequality for the youngest cohort 1981-1990, of which region of birth and father's occupation are the two key circumstances, accounting for 4.2 percent and 3.1 percent of the income inequality respectively. The contribution of ethnic minority status is relatively identical across birth cohorts, and is not as large as other circumstances, ranging from 1.0 percent to 1.6 percent.

5. The contribution of efforts in rural China

To examine the role of efforts other than circumstance in determining the total earnings inequality, we return to the method stated in Section 3.2. Besides the circumstances above, we identify four effort variables for further study. The first one is off-farm employment,

which indicates whether farmers choose to be off-farm employed. The share of respondents who report themselves involved in non-agricultural activities is 49 percent. Note that the ratio of migrants is only 23 percent, indicating that local employment offers an important avenue for farmers to be off-farm employed. The third effort variable we choose is education that has long been a key factor contributing to individual earnings in the literature. Given the rural context, we treat it as a dummy that indicates whether an individual has completed junior high school and above. Table 4 shows that more educated individuals who choose to migrate or be off-farm employed have higher earnings. The last "effort" variable included is marital status. The reason for including this variable is that couples, in contrast to single men or women, have interdependent preferences that affect their household income decisions (Zhang et al., 2008). Table 4 presents that married individuals are likely to earn less than those who are not married.

Table 4. Earnings and Efforts, by 5-year Cohorts

Birth Cohort	Share	Earnings			
		All	1951-1960	1961-1970	1971-1980	1981-1990
Education						
Below Junior High School	42	17554	14999	17575	21334	18011
Junior High and Above	58	35328	23352	35388	41461	35806
Off-farm employment						
No	51	15591	11999	16325	17966	20158
Yes	49	40626	35698	41292	44965	37657
Migration						
No	77	21914	13735	23965	25551	28577
Yes	23	26717	17987	27118	30638	26726
Marriage						
No	7	29425	17452	22439	25490	36507
Yes	93	27711	18136	27229	34138	32377

To further investigate how individual earnings is associated with circumstances and these 'effort' variables, we adopt Barry's scenario and Roemer's scenario using the approach of Jusot, Tubeufand Trannoy (2013) discussed in Section 3.2. In a first step, we show the impacts of circumstances on the four chosen effort variables using a Probit model (Table 5). Nearly all the coefficients take on their expected signs and relative magnitudes. For example, individuals with more educated fathers are more likely to be more educated themselves, and their decision of employment is determined by all the circumstances included: off-farm

employment is more likely to be the choice of male and Han individuals with better father's background. People born in the east are more likely to be off-farm employed as well, but less likely to migrate. This is in recognition of the fact that the east is the main destination for the majority of the migrants (Chan, 2013). Men are less likely to be married than women in rural China, which reflects the unbalanced sex ration and its impacts on rural men. Turning to the Pseudo R-square, the regressions of own-education and off-farm employment take on the highest values—0.18 and 0.19 respectively—while the migration regression takes on the lowest value (0.08), suggesting that the indirect effect of circumstances on earnings through migration is likely to be the smallest.

Table 5. The impacts of circumstances on efforts

	Effort Variables			
	Education	Migration	Off-farm	Marriage
Gender	0.599***	0.455***	0.336***	-0.291***
	(0.03)	(0.04)	(0.03)	(0.04)
Primary school	0.390***	0.205***	0.287***	-0.091*
	(0.03)	(0.04)	(0.03)	(0.05)
Junior high school	0.772***	0.146**	0.340***	-0.130*
	(0.05)	(0.06)	(0.05)	(0.07)
Senior high school	0.658***	0.100	0.397***	-0.154*
	(0.07)	(0.08)	(0.07)	(0.09)
Father-High status job	0.405***	0.176***	0.650***	-0.100*
	(0.04)	(0.05)	(0.04)	(0.05)
Farther- Han Chinese	0.450***	-0.092*	0.466***	0.005
	(0.05)	(0.05)	(0.05)	(0.07)
East	0.235***	-0.166***	0.601***	0.032
	(0.04)	(0.04)	(0.04)	(0.05)
Centre	0.126***	0.044	0.277***	0.038
	(0.04)	(0.05)	(0.04)	(0.06)
Birth cohort 61-70	0.410***	0.287***	0.505***	0.206***
	(0.04)	(0.05)	(0.04)	(0.06)
Birth cohort 71-80	0.546***	0.535***	0.899***	0.190***
	(0.04)	(0.05)	(0.04)	(0.07)
Birth cohort 81-90	1.155***	0.805***	1.213***	-0.727***
	(0.05)	(0.06)	(0.05)	(0.06)
Constant	-1.405***	-1.368***	-1.812***	1.822***
	(0.06)	(0.06)	(0.06)	(0.08)
N	9087	6966	9093	9096
Pseudo R2	0.179	0.073	0.187	0.112

Notes: *** 1%,** 5%,* 10%; Standard errors in parentheses.

Table 6 presents the regressions of individual earnings on both circumstances and effort variables shown in equation 9 and 10 for Roemer's scenario and Barry's scenario respectively. Column 1 and column 2 present the results when migration is excluded. All the coefficients have expected values and signs: married individuals with better education and

off-farm jobs are more likely to earn more. It is noteworthy that the magnitudes of all these effort variables are smaller in Roemer's scenario in which efforts are purged of any contamination coming from circumstances. For example, off-farm employment, in contrast to farm activities, is associated with 0.445 and 0.741 log points in Roemer's scenario and Barry's scenario respectively. The results in column 1 and 2 in which we exclude migration also show that the adjusted R-squared values increase considerably from 0.20 in Table 2 to 0.31 in Table 6.

Table 6. The impacts of circumstances and efforts on earnings

	Log(earnings)			
	Roemer	Barry	Roemer	Barry
Gender	0.481***	0.434***	0.486***	0.441***
	(0.0207)	(0.0211)	(0.0248)	(0.0254)
Primary school	0.158***	0.113***	0.130***	0.0906***
	(0.0241)	(0.0243)	(0.0279)	(0.0280)
Junior high school	0.191***	0.136***	0.140***	0.0909**
	(0.0357)	(0.0360)	(0.0443)	(0.0446)
Senior high school	0.244***	0.188***	0.192***	0.141**
	(0.0462)	(0.0464)	(0.0618)	(0.0618)
Father-High status job	0.0954***	0.0415	0.0363	-0.0140
	(0.0290)	(0.0292)	(0.0387)	(0.0389)
Farther- Han Chinese	0.171***	0.114***	0.163***	0.113***
	(0.0336)	(0.0337)	(0.0384)	(0.0386)
East	0.330***	0.270***	0.389***	0.332***
	(0.0258)	(0.0261)	(0.0308)	(0.0313)
Centre	0.230***	0.200***	0.251***	0.223***
	(0.0279)	(0.0279)	(0.0331)	(0.0332)
Birth cohort 61-70	0.375***	0.309***	0.399***	0.340***
	(0.0281)	(0.0283)	(0.0315)	(0.0318)
Birth cohort 71-80	0.559***	0.452***	0.584***	0.485***
	(0.0314)	(0.0320)	(0.0362)	(0.0369)
Birth cohort 81-90	0.532***	0.404***	0.555***	0.440***
	(0.0357)	(0.0369)	(0.0428)	(0.0445)
Education	0.163***	0.271***	0.140***	0.233***
	(0.0142)	(0.0237)	(0.0161)	(0.0269)
Off-farm employment	0.445***	0.741***	0.391***	0.653***
	(0.0142)	(0.0235)	(0.0170)	(0.0283)
Marriage	0.129***	0.221***	0.183***	0.319***
	(0.0230)	(0.0404)	(0.0278)	(0.0490)
Migration			0.0216	0.0345
			(0.0184)	(0.0298)
Constant	8.700***	8.040***	8.597***	7.902***
	(0.0413)	(0.0546)	(0.0488)	(0.0639)
Observations	9,084	9,084	6,957	6,957
R-squared	0.306	0.307	0.266	0.267

Notes: *** 1%,** 5%,* 10%; Standard errors in parentheses.

However, the last two columns in Table 6 show that the adjusted R-squared values turn out to be much lower when migration is included. This is mainly because there are much more missing values for migrants, resulting in the loss of more than 2000 observations. Most importantly, migration turns out to have an insignificant impact on individual earnings. Along with the results shown in Table 5—the low Pseudo R-square value, this indicates that circumstances influence individual earnings, in not only a direct way, but also an indirect way mainly through three of the four chosen effort variables—education, off-farm employment and marital status. Thus, besides own education and marital status, we mainly focus on off-farm employment rather than migration in the following analysis.

Using the estimated coefficients in Table 6, we then assess how the contributions of the circumstances and efforts change with two alternative views. Note that migration is excluded based on the discussion above. Roemer's view would maximise the magnitude of inequality in individual earnings due to circumstances and minimize the magnitude of inequality due to efforts. Table 7 shows that the contribution of circumstances to income inequality is 49.2 percent and 39.2 percent in Roemer and Barry's scenarios respectively. It is much larger than the figure in Section 4.3 (20.4 percent) for the full sample. As a note of caution, these two approaches are not directly comparable as we discussed above. But, bearing this in mind, we can at least conclude that only three efforts—off-farm employment, education and marriage—account for 60.8% of the income inequality in Barry's scenario, and the contribution of "effort" should be larger with the inclusion of other omitted effort variables that individuals may exert to influence individual earnings. Most importantly, the difference of the circumstantial contribution between Roemer's and Barry's scenario (10 percent) provides further evidence for the indirect component of the circumstantial inequality, which works through the 'effort' variables.

Table 7. Decomposition of inequalities according to both circumstances and efforts

Full model	Contribution of circumstances (%)	Contribution of efforts (%)	Total inequality
Roemer's scenario	49.2%	50.8%	0.41
Barry's scenario	39.2%	60.8%	0.41

6 Conclusions

This paper seeks to quantify the role of inequality of opportunity—associated with family background, gender, ethnic minority status, region of origin and birth cohort— in generating in individual earnings in rural China. Using the China Labour-force Dynamics Survey (CLDS) for 2014, this paper is the first attempt in a Chinese context, to the best of our knowledge, to examine whether and to what extent circumstantial inequality contributes to total income inequality.

The empirical results, based on the OLS regressions, show that the share of inequality of opportunity in individual earnings for the full sample is 20.4 percent of the GE(0) coefficient. The Shapley-value decomposition shows that besides gender, birth cohort and birth region which are dominant in contributing to the overall inequality of opportunity, father's background are playing an important role as well. This result varies across birth cohorts: the youngest cohort 1981-1990 faces the lowest total inequality in earnings, but it turns out to be the one with highest total circumstantial inequality as well as the partial inequality of opportunities stemming from each of the circumstances, with the only exception of gender. This verifies the recent popular saying—"daddy is the key" –in China, which, in a broader context, means that circumstantial inequality matters in economic outcome in rural China.

A closer investigation to include both circumstances and 'effort' in one framework indicates the importance of three effort variables including education, marital status and off-farm employment, which account for 60.8% of the earnings inequality. Circumstances influence individual earnings, not only directly, but also indirectly through these three effort variables. Most importantly, our result shows that migration is not playing a role.

The IOR in this study (20.4 percent) is smaller in contrast to those shown in Ferreira & Gignoux (2011) –32.2 percent in Brazil, 33.5 percent in Guatemala, 30.1 percent in Panama, 27.9 percent in Perun, 25.9 percent in Colombia—and some others, such as around 27 percent in Egypt (Assaad et al., 2017), but larger than the case of some high welfare states, such as Sweden shown in Björklund, Jänttiand Roemer (2011). However, it is extremely close to the case of India where overall opportunity share in total observed inequality is 20.8 percent for the rural sample (Singh, 2012). As China and India are the two largest developing nations in the world, the similarities between them are significant, especially for the rural areas. This gives us some confidence in this result. Moreover, for the following study, the inclusion of

both rural and urban samples is likely to yield a higher IOR because the *hukou* status as a circumstance is a great contributor to rural-urban income gap in China (Li & Gibson, 2013; Sicular et al., 2007; Yang, 1999).

Even with the above comparisons, inequality of opportunity in rural China is still worthy of attention as a substantial part (more than 20%) of total earnings inequality is accounted by unequal circumstances. But a question remains still on the table: whether inequality stemming from efforts, as well as from circumstances should be compensated? There is the philosophical belief, held by many, that individuals deserve to benefit from their inborn traits (Nozick, 1974). The equal-opportunity view, however, maintains that equality of opportunity is achieved when circumstances do not play any role in the resulting outcome (Roemer, 1998) and "economic inequalities due to factors beyond the individual responsibility are inequitable and to be compensated by society, whereas inequalities due to personal responsibility are equitable and not to be compensated" (Peragine, 2004). We side with the latter, and further argue that a substantial part of "effort" should not be considered voluntary, but is due to circumstances that lie beyond the control of individuals. For social policy, this is important, but difficult as well, with the requirement of understanding to what extent "efforts" are due to circumstances.

Although equality of opportunity is somewhat utopian and far from achieved, it does give the Chinese government an insight into two kinds of policies used for reducing inequality of opportunities in society: (i) policies focused on zero discrimination in opportunities and (ii) policies aimed at mitigating the impact of family background on child's chances of acquiring skills and abilities (Singh, 2012). In the case of rural China discussed in this study, it is of great significance for China's government to take affirmative action to reduce the inequality stems from gender, region of birth and father's background which turn out to be important contributors to individual earnings in both a direct and an indirect way working through efforts, such as off-farm employment and own education. Policies targeting these underlying problems are likely to reduce the overall earnings inequality since inequality of opportunity accounts for a significant part of total earnings inequality in rural China.

References

Arneson, R. J. (1989). Equality and Equal-Opportunity for Welfare. *Philosophical Studies, 56*(1), 77-93.

Assaad, R., Krafft, C., Roemer, J., & Salehi-Isfahani, D. (2017). Inequality of Opportunity in Wages and Consumption in Egypt. *Review of Income and Wealth*.

Barry, B. X. (2005). *Why social justice matters*: Polity.

Benjamin, D., Brandt, L., & Giles, J. (2005). The evolution of income inequality in rural China. *Economic Development and Cultural Change, 53*(4), 769-824.

Björklund, A., Jäntti, M., & Roemer, J. E. (2011). Equality of opportunity and the distribution of long-run income in Sweden. *Social Choice and Welfare, 39*(2-3), 675-696.

Bonnefond, C., & Clément, M. (2012). An analysis of income polarisation in rural and urban China. *Post-Communist Economies, 24*(1), 15-37.

Bourguignon, F., Ferreira, F. H., & Menendez, M. (2007). Inequality of opportunity in Brazil. *Review of income and Wealth, 53*(4), 585-618.

Bourguignon, F., Ferreira, F. H. G., & Menéndez, M. (2013). Inequality of Opportunity in Brazil: A Corrigendum. *Review of Income and Wealth, 59*(3), 551-555.

Bowles, S. (1972). Schooling and inequality from generation to generation. *Journal of Political Economy, 80*(3), S219-S251.

Chan, K. W. (2013). China: internal migration.

Checchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *The Journal of Economic Inequality, 8*(4), 429-450.

Checchi, D., Peragine, V., & Serlenga, L. (2010). Fair and unfair income inequalities in Europe.

Chen, Z., Ge, Y., Lai, H., & Wan, C. (2013). Globalization and Gender Wage Inequality in China. *World Development, 44*, 256-266.

Clément, M. (2016). Income mobility and income inequality in rural China. *Frontiers of Economics in China, 11*(4), 1673-3444.

Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics, 99*(4), 906-944

Dworkin, R. (1981). What is equality? Part 1: Equality of welfare. *Philosophy & public affairs*, 185-246.

Ferreira, F. H. G., & Gignoux, J. (2011). The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. *Review of Income and Wealth, 57*(4), 622-657.

Fleisher, B., Li, H., & Zhao, M. Q. (2010). Human capital, economic growth, and regional inequality in China. *Journal of Development Economics, 92*(2), 215-231.

Golley, J., & Kong, S. T. (2016). Inequality of opportunity in China's educational outcomes. *China Economic Review*.

Gourieroux, C., Monfort, A., Renault, E., & Trognon, A. (1987). Generalised residuals. *Journal of econometrics, 34*(1-2), 0304-4076.

Gustafsson, B., & Li, S. (2002). Income inequality within and across counties in rural China 1988 and 1995. *Journal of Development Economics, 69*(1), 179-204.

Gustafsson, B., & Shi, L. (2003). The ethnic minority-majority income gap in rural China during transition. *Economic Development and Cultural Change, 51*(4), 805-822

Hannum, E. (2005). Market Transition: Educational Disparities, and Family Strategies in Rural China: New Evidence on Gender Stratification and Development. *Demography, 42*(2), 275-299.

Hanoch, G. (1967). An economic analysis of earnings and schooling. *Journal of human Resources*, 310-329.

Howell, A. (2017). Impacts of Migration and Remittances on Ethnic Income Inequality in Rural China. *World Development, 94*(Migration on income), 200-211.

Juarez, F. W. C., & Soloaga, I. (2014). iop: Estimating ex-ante inequality of opportunity. *Stata Journal, 14*(4), 830-846.

Jusot, F., Tubeuf, S., & Trannoy, A. (2013). Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health? *Health economics, 22*(12), 1470-1495.

Kung, J. K., & Lee, Y.-f. (2001). So what if there is income inequality? The distributive consequence of nonfarm employment in rural China. *Economic Development and Cultural Change, 50*(1), 19-46.

Lefranc, A., Pistolesi, N., & Trannoy, A. (2008). Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike? *Review of Income and Wealth, 54*(4), 513-546.

Li, C., & Gibson, J. (2013). Rising Regional Inequality in China: Fact or Artifact? *World Development, 47*, 16-29.

Matthews, R., & Nee, V. (2000). Gender Inequality and Economic Growth in Rural China. *Social Science Research, 29*(4), 606-632.

Morduch, J., & Sicular, T. (2000). Politics, growth, and inequality in rural China: does it pay to join the Party? *Journal of Public Economics, 77*(3), 331-356.

Naughton, B. (2007). *The Chinese economy: Transitions and growth*: MIT press.

Nozick, R. (1974). *Anarchy, state, and utopia* (Vol. 5038 %@ 0465097200): New York: Basic Books.

Peragine, V. (2004). Ranking income distributions according to equality of opportunity. *Journal of Economic Inequality, 2*(1), 11-30 %@ 1569-1721.

Peragine, V., & Ferreira, F. (2015). Equality of Opportunity: Theory and Evidence. *World Bank Policy Research Paper*(7217).

Roemer, J. E. (1993). A pragmatic theory of responsibility for the egalitarian planner. *Philosophy & Public Affairs*, 146-166.

Roemer, J. E. (1998). *Equality of Opportunity* Harvard University Press, Cambridge, MA.

Scharf, M. M., & Rahut, D. B. (2014). Nonfarm Employment and Rural Welfare: Evidence from the Himalayas. *American Journal of Agricultural Economics, 96*(4), 1183-1197.

Shorrocks, A. F. (1982). Inequality decomposition by factor components. *Econometrica: Journal of the Econometric Society*, 193-211

Shorrocks, A. F. (2012). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *The Journal of Economic Inequality, 11*(1), 99-126.

Sicular, T., Ximing, Y., Gustafsson, B., & Shi, L. (2007). The urban–rural income gap and inequality in China. *Review of Income and Wealth, 53*(1), 93-126.

Singh, A. (2012). Inequality of Opportunity in Earnings and Consumption Expenditure: The Case of Indian Men. *Review of Income and Wealth, 58*(1), 79-106.

Walder, A. G. (2002). Markets and income inequality in rural China: Political advantage in an expanding economy. *American Sociological Review*, 231-253.

Wan, G. H., & Zhou, Z. Y. (2005). Income inequality in rural China: Regression-based decomposition using household data. *Review of Development Economics, 9*(1), 107-120.

Weiss, R. D. (1970). The effect of education on the earnings of blacks and whites. *The Review of Economics and Statistics*, 150-159.

Xie, Y., & Zhou, X. (2014). Income inequality in today's China. *Proc Natl Acad Sci U S A, 111*(19), 6928-6933.

Yang, D. T. (1999). Urban-biased policies and rising income inequality in China. *The American Economic Review, 89*(2), 306-310.

Zhang, J., Han, J., Liu, P.-W., & Zhao, Y. (2008). Trends in the gender earnings differential in urban China, 1988–2004. *ILR Review, 61*(2), 224-243.

Zhang, Y., & Eriksson, T. (2010). Inequality of opportunity and income inequality in nine Chinese provinces, 1989–2006. *China Economic Review, 21*(4), 607-616.

Zhou, Y., Han, H., & Harrell, S. (2008). From Labour to Capital: Intra-Village Inequality in Rural China, 1988–2006. *The China Quarterly, 195*.