



Pre-market discrimination or post-market discrimination: research on inequality of opportunity for labor income in China

Yue Yu¹ · Chengkui Liu¹

Received: 10 August 2021 / Accepted: 23 September 2022 / Published online: 6 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Inequality of opportunity (hereafter “IO”) restricts the realization of social justice, and its mechanism has always attracted attention. Using China Labor-force Dynamic Survey (CLDS) data in 2012, 2014, and 2016, we fully consider the impact of easily neglected educational opportunities on income inequality and creatively get pre-market and post-market discrimination channels. Research shows that IO is a fundamental cause of employees’ income inequality in China. Male–female and urban–rural opportunity inequality can severally explain 31.66% and 17.16% of total IO. Using the optimized Oaxaca–Blinder decomposition method, we obtain that pre-market and post-market discriminations are the main paths of urban–rural and male–female opportunity inequality, respectively. What’s more, the above pathway has different characteristics in different income groups. Therein, affected by the asymmetric information and employer prejudice in the labor market, the proportion of post-market discrimination channels shows a downward trend as income increases. The conclusions provide empirical support for eliminating market discrimination and ensuring equality of opportunity. They also have enlightening significance for relevant policy formulation.

Keywords Pre-market discrimination · Post-market discrimination · Inequality of opportunity · Signaling value of education · Employer prejudice

1 Introduction

The reality is that unbalanced and insufficient economic development has become the central contradiction in China. The unfair income distribution is still challenging to solve. In 2018, statistics show that China’s Gini coefficient of disposable income per

✉ Chengkui Liu
chamberlain23@163.com

¹ Economics and Management School, Wuhan University, Wuhan, China

capita still exceeds 0.4 (i.e., the international alert level). However, the Gini coefficient ignores the rationality of the sources of inequality. We readily accept inequality due to individual efforts but cannot tolerate inequality caused by circumstances such as gender, race, and other factors. In other words, individuals pay more attention to the sources of income inequality than to income inequality itself. An individual's perception of income inequality mainly depends on whether the source of income is fair (see, e.g., Cappelen et al. 2010).

Roemer's (1993, 1998) pioneering work on the inequality of opportunity (hereafter "IO") provides a new perspective for studying the sources of income inequality. All determinants of income, according to Roemer's paradigm, can be divided into two categories: "circumstances," which are out of an individual's control (for example, gender or birthplace), and "efforts," which are affected by individuals or not exogenous to the individual economically. Thus, income inequality caused by "circumstances" is IO, and income inequality due to "efforts" is inequality of effort.

Following Roemer's "circumstance–effort" dichotomy, academia has made multi-dimensional progress in exploring the factors affecting IO: On the one hand, focused "circumstance" factors include gender, hukou (the household registration system called hukou in Chinese), and family background largely explain the sources of inequality (see, e.g., Fuller and Vosko 2008; Björklund et al. 2012); on the other hand, focused "effort" factors include diploma, occupation, and others (see, e.g., Checchi and Peragine 2010; Ordine and Rose 2011). Moreover, some scholars are committed to improving methods and fully considering the interaction between circumstance and effort factors to reduce measurement errors (see Jusot et al. 2013).

The existing literature shows that the IO in developing countries is high, accounting for more than 30% of the total income gap. In contrast, the IO in developed countries is generally low, usually explaining less than 20% of the income gap (e.g., Lefranc et al. 2008; Ferreira and Gignoux 2011). Although research design and samples used affected the answer of how much disparity is attributable to opportunity, IO accounts for more than 35% of the income gap in most cases in China (see, e.g., Shi et al. 2018; Li and Lv 2019).

What's more, gender frequently remains the most prominent circumstance contributor in China. In Golley et al. (2019), gender explains 28% of IO. Based on different databases, Yang et al. (2021) reported that its ratio in IO has increased significantly over time, from 31.63% in 2013 to 36.9% in 2018. According to Li and Lv (2019), the contribution of hukou to IO is 17.51%, 12.42%, 12.8%, and 10.81% in 2008, 2010, 2013, and 2015, respectively.

However, how is the IO formed? Actually, IO occurs both before and after individuals enter the labor market and violates "the level-the-playing-field principle" and "the nondiscrimination principle," respectively (see Roemer 1998). When the former is also considered a kind of discrimination, we can claim that the IO results from pre-market and post-market discrimination. Indeed, the above existing studies did not distinguish the values pre- and post-market.

On the one hand, post-market discrimination is mainly embodied in "different pay for equal work." There is some consensus on the key factors causing IO in the existing literature, "circumstance" factors discriminated against principally include gender

and hukou (see, e.g., Meng and Zhang 2001; Kübler et al. 2018). As for male–female inequality, the International Labor Organization (ILO) “2018–2019 Global Wage Report” shows that this discrimination is widespread worldwide, and women’s average monthly income is only 80% of men’s.¹ As for urban–rural inequality, China’s dual labor market caused by the particular hukou system makes it inevitable that urban–rural opportunity inequality in wage premiums and employment (see, e.g., Zhang et al. 2016). Besides, the urban- or rural-born will also significantly impact employees’ occupational mobility (see, e.g., Liu 2012; Yamamoto et al. 2019).

On the other hand, as in “the level-the-playing-field principle,” equality of opportunity does not only mean no discrimination during the competition if work is like a competitive game. However, it also must ensure that everyone has to step into the same “arena” before competing. The same “arena,” based on the life cycle of “education before work,” means that educational opportunities for all should not differ by “circumstances” (see Roemer 1998; Démurger et al. 2009).

Therefore, the pre-market mechanism for IO is the indirect effect in light of the fact that the circumstances create differences in human capital through educational opportunities, which further leads to income inequality in the labor market (Neal and Johnson 1996; Krafft and Alawode 2018). As for male–female inequality especially, there is a common occurrence of “son preference” in rural China that parents will prioritize their sons in education, even if the daughter is hardworking. Then, how big is the effect of pre-market discrimination caused by circumstances? According to Chetty et al. (2014), based on 216 different US commuting zones, individual educational skills can explain 12% of the variation in the intergenerational income elasticity. What’s more, Palomino et al. (2019) show that education on pre-market accounts for up to 30% of total IO in 26 European countries.

In conclusion, we regard the path in which “circumstances” indirectly affect income through educational opportunities, before one’s labor force participation, as the pre-market discrimination channel of IO. The post-market discrimination channel, by comparison, is how “circumstances” directly affect income after individuals enter the labor market.

Generally, this paper has the following main contributions: (1) Existing research does not pay enough attention to the mechanism of IO. Hence we creatively obtain the pre-market and post-market discrimination pathways of IO to measure the impact of IO more comprehensively, which is rare in existing studies. What’s more, we also compare the above channels in male–female and urban–rural opportunity inequality. (2) In methods, we contribute to extending the Oaxaca–Blinder decomposition to estimate specific values of the above channels on IO with China data. (3) Existing research focuses on measuring the value of IO or factors decomposition but lacks theoretical support. Therefore, we try to show pre-market and post-market discrimination further in the formation of IO through the theoretical model based on the discrimination theory and asymmetric information.

¹ “Global Wage Report 2018/19,” ILO official Web site, 2018–11–26, <http://www.ilo.org/global/research/global-reports/global-wage-report/2018/lang-en/index.htm>.

2 The model

As mentioned above, we see the impact of “circumstances” on educational opportunities, before one’s labor force participation, as the pre-market discrimination channel in IO. In contrast, the post-market discrimination path is how “circumstances” directly affect one’s income in the labor market.

2.1 Pre-market discrimination

We devote this section to demonstrating pre-market discrimination against “circumstances.” Taking the male–female inequality in educational opportunities as an example, we propose the utility function for family, written as

$$U = C_t + \varepsilon [\Gamma(E_{m,t+1}, E_{f,t+1}) - \delta_m E_{m,t+1} - \delta_f E_{f,t+1}] - \gamma E_{f,t+1} \quad (1)$$

where C_t is household consumption in the current period t , ε is the discount factor, and Γ is expected return to education, which takes on two parameters: household expected years of schooling for their children that are boys $E_{m,t+1}$ and girls $E_{f,t+1}$. Let δ_m and δ_f denote the marginal cost of education for boys and girls, respectively. γ is the degree of “son preference” in the household, and then a larger γ indicates the more likely the household will support boys’ studies rather than girls.

For simplicity, we assume that Γ satisfies Cobb–Douglas form, that is

$$\Gamma(E_{m,t+1}, E_{f,t+1}) = E_{m,t+1}^u E_{f,t+1}^v, \text{ with } (0 < u, v < 1) \quad (2)$$

where both u and v are expected return rates on education. The household expectation can affect children’s educational opportunities; a higher expectation implies that the household is more likely to increase investment in children’s education (refer to Sewell and Shah 1968; Spera et al. 2008; Heckman and Kautz 2013). Thus, we use the ratio of household expectations to illustrate the gender gap in access to education (i.e., pre-market discrimination). Using the first-order condition, one can derive

$$E_{f,t+1}/E_{m,t+1} = (v/u) \times [\varepsilon\delta_m/(\varepsilon\delta_f + \gamma)] \quad (3)$$

where $(E_{f,t+1}/E_{m,t+1}) < 1$ implies the disadvantages of females in educational opportunities. Moreover, we assume that both δ_m and δ_f are constant in time and common to all households. Then, Eq. (3) shows that pre-market discrimination against females is the result of statistical discrimination (i.e., the gender difference in expected return rate, $v < u$) or taste-based discrimination in the household (i.e., the patriarchal concept, $\gamma > 0$).

2.2 Post-market discrimination

This section builds on the statistical literature (see Arrow 1973) and illustrates post-market discrimination against “circumstances.” We assume, taking the gender in the “circumstances” as an example again, that the enterprise’s output is

$$Y = \lambda(Q^m + Q^f) - g(S^m + S^f) \tag{4}$$

where i is individual gender, $i = m$ indicates the male, $i = f$ indicates the female. λ is the qualified individual’s productivity, with $\lambda \in (0, 1)$, whereas the productivity of unqualified individuals is zero. g is the unit production cost and satisfies $g > 0$.

Considering existing research on discrimination in the labor market (see Coate and Loury 1993; Conde-Ruiz et al. 2022), we assume that employers observe each individual’s “circumstances” identification (“gender” in here) and signals (e.g., test scores, work experience) to determine whether the individual is qualified and the individual’s wage, given the individual with a higher signal value has the more significant posterior likelihood that he is qualified. Let ϵ ($0 \leq \epsilon \leq 1$) and $\hat{\epsilon}^i$ be the signal sent by individuals and the hiring threshold. Thus, employed individuals meet $\epsilon \geq \hat{\epsilon}^i$. Then, one can derive the number of employed S^i and qualified individuals Q^i , written as

$$S^i = \left[\theta^i (1 - F_b(\hat{\epsilon}^i)) + (1 - \theta^i) (1 - F_d(\hat{\epsilon}^i)) \right] \bar{S}^i \tag{5}$$

$$Q^i = \theta^i (1 - F_b(\hat{\epsilon}^i)) \bar{S}^i \tag{6}$$

where θ^i is the employer’s belief for the proportion of qualified individuals, \bar{S}^i is the total number of job seekers, $F_b(\epsilon)$ ($F_d(\epsilon)$) is the cumulative distribution function of qualified (unqualified) individuals, and $f_b(\epsilon)$ ($f_d(\epsilon)$) is the corresponding density function. We, in addition, assume that $f_b(\epsilon)/f_d(\epsilon)$ is non-increasing on $[0, 1]$. Using the first-order condition, the wage in light of signals, $I^i(\epsilon)$, is given by

$$I^i(\epsilon) = [\lambda \theta f_b(\epsilon) / (\theta f_b(\epsilon) + (1 - \theta) f_d(\epsilon))] - g \tag{7}$$

Let $\Lambda(\theta, \epsilon) = \theta f_b(\epsilon) / [\theta f_b(\epsilon) + (1 - \theta) f_d(\epsilon)]$, then we can reach the following:

$$I^m(\epsilon) - I^f(\epsilon) = \lambda \left[\Lambda(\theta^m, \epsilon) - \Lambda(\theta^f, \epsilon) \right] \tag{8}$$

Equation (8) shows that given the same productivity and signals after individuals join the labor market, the inequality comes from post-market discrimination against gender due to employer’s prejudice (θ^m and θ^f) under incomplete information (indicated by ϵ).

3 Methodology

3.1 Inequality of opportunity

In Roemer's framework (see Roemer 1993, 1998), all determinants of income can be partitioned into two components: "circumstances" (denoted by C_i) and "efforts" (denoted by E_i). Specifically, "circumstances" are beyond individual control (e.g., gender, birthplace, or race), and "efforts" are not economically exogenous to the individual.

In the sense of Roemer, equality of opportunity would be achieved when "circumstances" have no impact on income. Then, individuals can be divided into T exhaustive and mutually exclusive types, where individuals in the same type have the same "circumstances." Let $F(Y|C^t)$ denote the income distribution for individuals of type t , where $t = 1, \dots, T$. Equality of opportunity demands that the income is independent of "circumstances," written as

$$F(Y) = F(Y|C^t) = F(Y|C^l) \quad (9)$$

where $t, l = 1, \dots, T$. However, testing whether the above condition holds is difficult. This article adopts the ex ante parametric method since there is no limit on the number of "circumstance" variables (see Ferreira and Gignoux 2011; Davillas and Jones 2020). Assuming income is a function of "circumstances" and "efforts," we estimate the following log-linearized equation:

$$\ln Y_i = \alpha + \beta C_i + \gamma E_i + \mu_i \quad (10)$$

where Y_i is the individual income, α is a constant term, C_i is the matrix of an individual's "circumstances," E_i is the matrix of an individual's "efforts," and μ_i is the residual term, which contains unobservable factors such as luck.

To get the counterfactual distribution, we, consistent with most literature works, treat "efforts" as the function of "circumstances"² (see, e.g., Bourguignon et al. 2007; Ferreira and Gignoux 2011), as shown in Eq. (11).

$$E_i = \omega + \lambda C_i + v_i \quad (11)$$

For measuring IO—instead of estimating any causal relationship between income, "circumstances," and "efforts"—we can write the reduced form by substituting Eq. (11) into Eq. (10), as follows:

$$\ln Y_i = \delta + \psi C_i + \varepsilon_i \quad (12)$$

² As for "effort" variables affected by "circumstance" variables in most literature works, evidence is that the individual's level of schooling is widely considered influenced by their family background (see, for example, Piketty 1995). In this regard, Roemer (1998) also takes the example, Asian students tend to study harder than others in the USA, to demonstrate that "individual 'efforts' are the reflection of 'circumstances'."

where $\delta = \alpha + \gamma\omega$, $\psi = \beta + \gamma\lambda$, $\varepsilon_i = \gamma v_i + \mu_i$.

Then, the counterfactual distribution with the interference of “efforts” removed is

$$\hat{Y}_i = \exp(\hat{\delta} + \hat{\psi} C_i) \quad (13)$$

where \hat{Y}_i is the circumstance-conditioned income, $\hat{\delta}$ and $\hat{\psi}$ are the estimates of δ and ψ in Eq. (12), respectively. Then, $\hat{\psi}$ includes the direct and indirect effects of “circumstances.”³ Let IO represent the inequality indicator (MLD index). IO—inequality between groups—is computed by $\text{IO}(\hat{Y}_i)$. Meanwhile, relative IO is $\text{IO}_r = \text{IO}(\hat{Y}_i)/\text{IO}(Y_i)$.

Besides, due to unobservable factors, the above result about inequality is the actual value’s lower bound. Considering that the MLD index is sensitive to low-income samples, this paper also reports the Gini coefficient (sensitive to middle-income samples) and Theil index (sensitive to high-income samples) to ensure the robustness of the results.

This paper uses the Shapley value method to explore the causes and differences of inequality based on the above measurement result. The ideas are as follows:

Firstly, we would exclude anyone “circumstance” variable from the income equation. Secondly, we would estimate the income equation again to obtain a new counterfactual distribution and IO result. At this time, the result no longer includes the influence of the eliminated “circumstance” variable. Finally, differences between results before and after elimination are the contribution of this “circumstance” factor to the IO. Since the contribution value is related to the order in which it is eliminated, this paper uses the average contribution value in each order.

3.2 Discrimination in pre-market and post-market

As in the level-the-playing-field principle proposed by Roemer (1993, 1998), equality of opportunity means no discrimination before and after. However, general decomposition methods cannot include pre-market discrimination. We build on Fleurbaey and Schokkaert (2011) and contribute to extending the Oaxaca–Blinder approach to decompose the pre-market and post-market discrimination channel simultaneously in measuring equality of opportunity. The process is as follows:

For ease of exposition, consider exploring the male–female opportunity inequality (the urban–rural inequality is the same and repeats no more). Suppose all individuals

³ Unavoidable omissions have not virtually affected the conclusion, as follows: (1) If unobserved “circumstance” or “effort” variables are uncorrelated with variables in Eq. (12), then the estimated coefficient is unbiased. Moreover, measured IO is still a lower bound of the actual, and omitted “efforts” do not affect the measured value. (2) If unobserved “circumstances” are correlated with variables in Eq. (12), the estimated coefficient has captured the contribution of omitted “circumstances,” measured IO is higher than the observed value but closer to the actual. (3) If unobserved “effort” variables are correlated with observed “circumstances” in Eq. (12), the coefficient has incorporated the indirect effect of “circumstances” on omitted “efforts,” measured value is closer to the actual.

have a commonly distributed error term, σ_i , and then we propose the measure of pre-market discrimination based on (10), for individual i , written as

$$\ln Y_i = \alpha + \beta C_i^m + (\tau + \phi C_i^m) C_i^n + (\rho + \xi C_i^m) E_i^m + (\zeta + \varphi C_i^m) E_i^n + \sigma_i \quad (14)$$

where C_i^m is a dummy indicating the group of individual i , i.e., $C_i^m = 1$ if individual i is male, $C_i^m = 0$ if individual i is female. Let E_i^m be individual years of schooling, and let C_i^n and E_i^n denote other uncontrollable and controllable factors (namely, “circumstances” and “efforts”), respectively.

According to Roemer (1993, 1998), individuals’ income should depend on their relative efforts. To express the relative effort as functions of observable variables in light of previous research works (see, e.g., Checchi and Peragine 2010), define:

$$R_i = \Gamma(\Delta^S - \Delta^F) \quad (15)$$

where R_i is the individual’s relative effort, Δ^S is the difference in years of schooling between individuals and their peers and Δ^F is the difference in years of schooling between the individual’s father and his father’s peers. Indeed, Δ^S indicates the excess effort of individuals and Δ^F indicates the excess support of the circumstances. Let quantile function Γ convert the relative effort into the group’s quantile.

If there is no pre-market discrimination, the relative effort rather than gender determines individual years of schooling. Then, the individual’s relative effort can be measured by

$$E_i^m = \Psi(C_i^m, R_i) = \eta + \theta C_i^m + (\vartheta + \iota C_i^m) R_i + \epsilon_i \quad (16)$$

Substitute (15) into (16) and estimate simultaneous formulas consisting of (14) and (16) via 3SLS (three-stage least square) to get IO consisting of pre-market and post-market discrimination, denoted by D:

$$D = \hat{\beta} + \hat{\varphi} K(E_i^n|f) + \hat{\phi} K(C_i^n|f) + \hat{\xi} K[\Psi(m, R_i|f)] + \hat{\rho} K[\Psi(m, R_i|f)] - \hat{\rho} K(E_i^m|f) \quad (17)$$

Among them, $\hat{\beta}$, $\hat{\varphi}$, $\hat{\xi}$, $\hat{\phi}$, and $\hat{\rho}$ are the estimated values and K indicates sample means, i.e., $K(E_i^n|f)$ is the sample mean of E_i^n using only observations where $C_i^m = 0$ (individual i is female). In (17), items 1 to 4 are differences in coefficients resulting from post-market discrimination.

It is worth noting that the last two terms, in (17), are IO attributable to pre-market discrimination. Under the premise of the same effort and return on education, the difference between the two items represents the resultant difference in years of schooling caused by unequal access to education before individuals enter the labor market.

4 Data

The research data in this paper are obtained from the China Labor-force Dynamic Survey (CLDS).⁴ Carried out by Sun Yat-Sen University, the survey is extensive in scale, broad in coverage, and highly targeted, including three levels of labor individuals, families, and communities. It is a comprehensive database with the labor population aged 15–64 as the primary survey object. Given the applicability of the improved decomposition method used in this paper, we processed CLDS into cross section data composed of all samples in CLDS2012 (it is the baseline survey) and new samples in CLDS2014 and CLDS2016.⁵ Then, all monetary variables are deflated by the GDP deflator. Following the principles of data rationality and availability, the selected variables are as follows:

4.1 Income

Considering that discrimination is more closely related to wage, we choose the employee's wage as a measurement indicator.⁶ The data comes from the questionnaire question, "how much is your wage?" We also take the logarithm of income and perform 1% tailing to obtain a more stable distribution result.

4.2 Circumstances

On the one hand, personal characteristics variables include an employee's age, ethnicity, gender, area, and hukou. The specific processing is as follows: (1) The equation includes the age and its quadratic term in the basic regression equation to explore the relationship between income and age. (2) We set ethnicity, gender, and area as dummy variables. Separately, the Han takes the value "1," and other ethnic groups take the value "0. Moreover, "1" is the male and "0" is the female. The east province takes the value "1," and the central and west provinces take the value "0." (3) The value of

⁴ The opinions are the author's alone. Please refer to <http://css.sysu.edu.cn> for more information about the CLDS data.

⁵ This processing is based on the following considerations: (1) The improved Oaxaca–Blinder method cannot decompose the values of inequality derived from the panel data, and the decomposition is almost all used for cross section data. (2) The key "circumstance" variable such as gender is time-invariant. We cannot estimate its effect when controlling individual fixed effects based on panel data, making it impossible to assess the exact counterfactual distribution to measure IO. (3) Representative research in measuring IO is almost all based on cross-section data (see, e.g., Bourguignon et al. 2007; Ferreira and Gignoux 2011; Palomino et al. 2019) or processes panel data into section data (see, e.g., Marrero and Rodríguez 2013). Thus, we also do this to guarantee that our conclusions are comparable.

⁶ We take the employee as our research object for the following reasons: One of our primary purposes is to obtain the pre-market and post-market discrimination channels by decomposition method and compare their contributions to IO. However, the premise of getting the results of post-market discrimination by definition is that individuals have the job-seeking need when they enter the job market, or there would be no such thing as post-market discrimination. In other words, employers, self-employed, or individuals who are not in the job market are not exposed to post-market discrimination. It makes no sense for us to explore the discrimination against groups that are not subject to it, and of course, our conclusions are only for the employee group.

urban hukou is “1,” while rural hukou value is “0,” and the data are obtained from the questionnaire question, “what was hukou at your birth?”

On the other hand, circumstances also include family background variables. Considering that fathers have more decision-making power in the family (see, e.g., Solon 1999), this paper uses the father’s relevant characteristics to illustrate family backgrounds. Their father’s occupation, education, and communist membership are selected variables. The specific treatment is as follows: (1) The father’s occupation at the age of 14 is chosen as an indicator to measure the family’s economic situation. This paper divides the father’s occupation type into seven levels based on the Chinese social occupation prestige score calculated by Li (2005) and assigns the value of 1–7 from low to high. The data are obtained from the questionnaire questions “father’s occupation at 14?” and “father’s work type at 14?” (2) The data on education are obtained from the questionnaire question, “What is your father’s education” and we convert it to the education years.⁷ (3) The father’s communist membership is set as a dummy variable: “1” is a communist and “0” is others.

4.3 Efforts

Effort variables include an employee’s education, communist membership, work attitude, marriage, and subjective health. The specific processing is as follows: (1) We convert everyone’s diploma into education years, and the data are obtained from the questionnaire question “what is your highest degree.” (2) The communist membership is set as a dummy variable: “1” is a communist and “0” is others. (3) We assign work attitude as a value of 1–4 from low to high according to the questionnaire question, “you agree that even if there are other reasons to take a break, I will work hard to complete what I should do every day.” The higher the willingness, the greater the value. (4) The marital status is set as a dummy variable: “1” is married and “0” is others. (5) We obtain the subjective health status from the questionnaire question “what do you think of your current health” and assign the value of 1–5 according to the health degree from ill to good.⁸

In addition, we also exclude samples with missing values and outliers in critical variables and exclude samples with employees younger than 16 years old or older than 60 years old. Descriptive statistics are shown in Table 1.

⁷ Specific assignments are: the number of education years for illiteracy is 0; the number of education years for elementary school or private school is 6; the number of education years for junior high school is 9; the value of high school, vocational high school, technical school, and technical secondary school is 12; the number of education years for a college degree is 15; 16 for bachelor’s degree; 19 for master’s degree and above.

⁸ We put health into the income equation and take it as the “effort” by definition for the following reasons. (1) Health is an essential part of human capital and undoubtedly affects personal income (see, e.g., French 2012). (2) The health of individuals is mostly under their control, such as through good living habits, etc. (see, e.g., Rosa Dias 2009; Donni et al. 2014; Fajardo-Gonzalez 2016). According to Rosa Dias (2009), about 79% of health inequality for British adults is due to individuals’ control factors. Donni et al. (2014) also estimate that controllable factors account for around 70%.

Table 1 Descriptive statistics for selected variables

Variable	Meaning of variables	Mean	SD	Median	Number of observations
Income	Individual income (Unit: 100 CNY)	3.409	0.776	3.401	2100
Age	Individual age	36.427	10.113	36	2248
Ethnic	China's ethnic groups	0.928	0.258	1	2248
Gender	Male or female	0.55	0.498	1	2248
Area	Central and west or east province	0.533	0.499	1	2248
Hukou	China's special household registration	0.338	0.473	0	2248
Occupation	Father's occupational prestige	2.713	1.242	2	2248
Diploma	Father's years of schooling	7.887	4.082	9	2244
Communist	Father's communist membership	0.176	0.381	0	2248
Education	Individual years of schooling	12.060	3.457	12	2248
Politics	Individual communist membership	0.223	0.417	0	1246
Attitude	Individual working attitude	3.019	0.574	3	1246
Marriage	Individual marital status	0.902	0.297	1	1246
Health	Individual subjective health	3.927	0.831	4	1246

5 Results

5.1 The measurement and influencing factors in IO

How much does IO explain income inequality? How much is the contribution of circumstances, and what is the specific mechanism of IO? Based on the estimated results of Eq. (13), this paper will describe in turn. Besides the MLD index, this paper also reports the Theil index and Gini coefficient results.

From Table 2, IO is a fundamental cause of the income gap. The overall degree is 0.037, accounting for 14.45% of the income gap. From the subsample results, different groups' characteristics in IO are different: on the one hand, females face higher IO. The IO can explain 16.27% of the male–female income gap, indicating that females need to overcome greater inequality. On the other hand, IO has roughly the same degree in urban- and rural-born, and explanatory power for the income gap is similar.

Table 2 Income inequality and inequality of opportunity

	All	Gender		Region		Province	
		Male	Female	Urban	Rural	East	Central and West
<i>Income inequality</i>							
MLD Index	0.256	0.218	0.295	0.242	0.249	0.266	0.241
Theil Index	0.244	0.210	0.285	0.227	0.239	0.255	0.226
Gini Index	0.413	0.382	0.446	0.393	0.413	0.418	0.404
<i>Inequality of opportunity</i>							
MLD Index	0.037	0.019	0.048	0.027	0.030	0.044	0.032
Theil Index	0.038	0.019	0.049	0.027	0.029	0.045	0.032
Gini Index	0.154	0.108	0.173	0.128	0.138	0.168	0.143
<i>Relative value</i>							
MLD Index	14.45%	8.72%	16.27%	11.16%	12.05%	16.54%	13.28%
Theil Index	15.57%	9.05%	17.19%	11.89%	12.13%	17.65%	14.16%
Gini Index	37.29%	28.27%	38.79%	32.57%	33.41%	40.19%	35.40%

Results show that the government has effectively balanced regional development and broken the urban–rural dual structure in recent years.

This paper further uses the Shapley value to decompose each circumstance's contribution value in the IO to judge the relative importance of each circumstance. From Table 3, we can get that the male–female opportunity inequality has an enormous contribution to the overall IO, accounting for 31.66%, and urban–rural opportunity inequality accounts for 17.16%. What's more, IO is also profoundly affected by family background, with a total contribution of 35.08%. Among them, the diploma accounts for the most significant proportion.

5.2 Decomposition

The above results show that the urban–rural and male–female opportunity inequality is the fundamental cause of the overall income gap. What is the influencing mechanism of circumstances individuals cannot control on the IO? According to Roemer (1998), the IO is explained by pre-market and post-market discrimination channels. The results are shown in Table 4 by estimating Eq. (17).

On the one hand, pre-market discrimination is the primary channel of urban–rural opportunity inequality. In other words, there is a more significant urban–rural difference in access to education. Hence the urban–rural opportunity inequality mainly comes from the effect of education inequality. As for specific values, the pre-market discrimination accounted for 79.2% of IO, while the post-market discrimination accounted for only 20.8%.

On the other hand, post-market discrimination is the central channel of male–female opportunity inequality. Namely, gender factors have a more significant impact on

Table 3 Contribution of circumstances based on Shapley value decomposition

	Personal characteristics				Province (%)	Family background		
	Gender (%)	Age (%)	Ethnic Groups (%)	Hukou (%)		Diploma (%)	Occupation (%)	Communist membership (%)
Full sample	31.66	5.87	5.56	17.16	4.66	18.15	12.31	4.62
Subsample								
Male	-	15.54	6.06	14.78	3.03	32.54	11.01	17.04
Female	-	5.95	8.46	32.41	8.14	20.30	22.52	2.22
Urban	14.75	7.56	6.45	-	22.36	16.19	24.58	8.12
Rural	63.05	7.89	7.51	-	2.52	15.94	1.27	1.83
East Province	22.16	9.07	0.98	22.57	-	20.03	21.57	3.61
Central and West	45.93	2.58	9.92	11.87	-	15.15	7.26	7.29

The individual's family background is set according to their father's characteristics

Table 4 Influencing channels of IO: pre-market and post-market discrimination

	Urban–rural			Male–female		
	Absolute value (1)	Relative value (2)	Robustness (3)	Absolute value (4)	Relative value (5)	Robustness (6)
<i>Full sample</i>						
Total value	0.269	1.000	– 0.252	0.196	1.000	– 0.181
Pre-market	0.213	0.792	– 0.096	– 0.027	– 0.138	0.017
Post-market	0.056	0.208	– 0.156	0.223	1.138	– 0.198
<i>Subsample: East</i>						
Total value	0.341	1.000	– 0.292	0.156	1.000	– 0.133
Pre-market	0.242	0.710	– 0.322	– 0.048	– 0.308	0.038
Post-market	0.099	0.290	0.030	0.204	1.308	– 0.171
<i>Subsample: Central and Western</i>						
Total value	0.202	1.000	– 0.223	0.241	1.000	– 0.224
Pre-market	0.172	0.851	0.098	– 0.010	– 0.041	0.006
Post-market	0.030	0.149	– 0.321	0.251	1.041	– 0.230
<i>Subsample: Ages [16,30]</i>						
Total value	0.390	1.000	– 0.311	0.248	1.000	– 0.257
Pre-market	0.224	0.574	– 0.007	– 0.008	– 0.032	0.007
Post-market	0.166	0.426	– 0.304	0.256	1.032	– 0.264
<i>Subsample: Ages [30,40]</i>						
Total value	0.406	1.000	– 0.353	0.173	1.000	– 0.228
Pre-market	0.206	0.507	– 0.207	– 0.051	– 0.295	0.014
Post-market	0.200	0.493	– 0.146	0.224	1.295	– 0.242
<i>Subsample: Ages [40,50]</i>						
Total value	0.329	1.000	– 0.311	0.187	1.000	– 0.114
Pre-market	0.352	1.070	– 0.196	0.052	0.278	– 0.037
Post-market	– 0.023	– 0.070	– 0.115	0.135	0.722	– 0.077
<i>Subsample: Ages [50,60]</i>						
Total value	0.053	1.000	– 0.016	0.11	1.000	0.475
Pre-market	0.066	1.245	0.280	0.002	0.018	0.003
Post-market	– 0.013	– 0.245	– 0.296	0.108	0.982	0.472

discrimination in the labor market than on the process of human capital accumulation. Results illustrate that the male–female difference in education acquisition has not been evident, but women are still in a weak position that is vulnerable to employer discrimination. To a certain extent, the above analysis also explains why women’s average education years are higher than men’s, but their income is lower than men’s.

As for subsample results, the primary influence channels have not changed in the east, central, and west regions. The pre-market discrimination always is the main

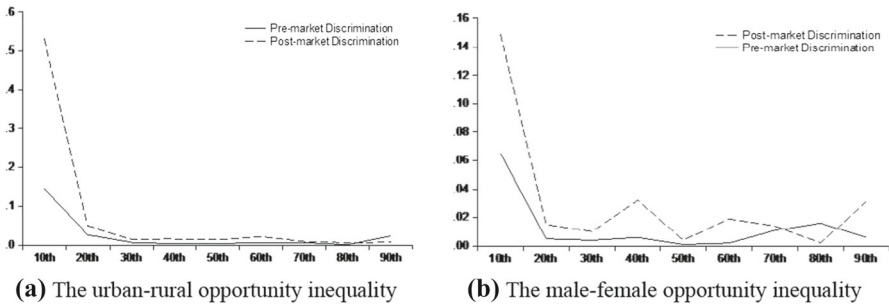


Fig. 1 Mechanism's characteristics at different income quantiles

pathway of the urban–rural opportunity inequality, and its proportions in the east and the central and west regions are 71.0% and 85.1%, respectively. For another, post-market discrimination is still the principal mechanism of male–female opportunity inequality.

According to the age-group, the proportion of pre-market discrimination in explaining urban–rural opportunity inequality has weakened over time, indicating that the new generation of employees will be more equitable in access to education. Post-market discrimination has mainly realized male–female opportunity inequality, but its proportion has gradually increased. Discrimination has increasingly become a constraint on improving the male–female income gap. Besides, considering that the decomposition results are affected by the choice of benchmarks, we replace the previous benchmark with urban and male samples, respectively, as shown in columns (3) and (6). By the results, the above conclusion is still valid.

Given that the above decomposition is based on the average income, we further explore the mechanism's characteristics at different income quantiles shown in Fig. 1. The x-axis represents income quantiles, and the y-axis represents pre-market and post-market discrimination.

Figure 1a shows that IO is on a downward trend with increased income. Separately, in the low-income group, the post-market discrimination can largely explain IO, which comes after individuals enter the labor market, such as employer prejudice. In contrast, IO gradually manifested as pre-market discrimination in the high-income group, formed in the accumulation stage of human capital.

Figure 1b shows that the change in the mechanism of male–female opportunity inequality is roughly the same as urban–rural inequality. In the low-income group, the post-market discrimination channel accounts for a more significant proportion. With the increase in income, the restrictions on the accumulation of human capital employees face are beginning to appear. Even in some higher-income groups, pre-market discrimination becomes the primary effect pathway.

6 Further discussion

The post-market discrimination is almost always the primary channel, whether it is urban–rural or male–female opportunity inequality. In addition, its proportion has gradually weakened as income increases. Then, what is the reason? We attempt to explain the phenomenon with information asymmetry and employer prejudice mentioned above.

6.1 Information asymmetry

According to the screening theory proposed by Spence (1973), employers cannot fully know an individual's actual productivity due to asymmetric information. Then, diplomas have the function of sending signals of productivity. Given that the larger the information asymmetry individuals face, the higher the signaling value of education, we can use the signaling value to illustrate the degree of information asymmetry (see Lang and Manove 2011).

In light of data features, the approach we used to obtain signaling value is estimating the sheepskin effect and is a well-established method in existing research (see, e.g., Ferrer and Riddell 2002; Schady 2003; Olfindo 2018). Precisely, we can estimate the extra benefits brought by getting the “sheepskin” (i.e., the diploma) to employees when controlling for their years of schooling (see, e.g., Jaeger and Page 1996; Bauer et al. 2005; Mora and Muro 2008).

Specifically, the coefficient of the diploma variable represents the rate of return on signals, which is the sheepskin effect and reflects the information function of education. The coefficient of years of schooling, by comparison, represents the net return on each additional year of education after deducting the diploma effect, which belongs to the human capital effect and reflects the productive function of education. Moreover, we also ruled out the possible case of multicollinearity in the model.

Table 5 shows that the net educational return rate in rural-born is 2.9%, less than 10.4% of urban-born, while the signaling value of a diploma at all stages is higher than that in urban-born. Specifically, the signaling values of the college, bachelor's, and master's degrees are 38.4%, 84.4%, and 54.0%,⁹ and all are significant at the 5% level. It shows that employers face a more significant information asymmetry regarding the rural-born. Then, there is a greater need for employees to send abilities signals to employers through diplomas, and the signaling value of education will be more excellent. Given the neglect of samples with diplomas but who have not entered the labor market, we use the Heckman model to deal with the sample selection bias and use the number of employees' children as the identification variable.¹⁰ Although

⁹ Generally speaking, employees who have obtained a bachelor's degree also have a high school degree, that is, the diploma effect has an accumulation effect. To avoid double counting, when calculating the diploma effect of a certain stage of education, it is necessary to exclude the influence of the diploma effect of the previous stage of education. Take the master's degree as an example, its diploma effect is $e^{(1.099-0.667)} - 1$.

¹⁰ See Wooldridge (2000, Chapter 17), and we need at least one identification variable that influences selection but not the overall model's dependent variable. The dependent variables of the overall model and the selection equation in this paper are individual wage and whether the individual chooses to work, respectively. We think that the number of children of individuals will affect their labor force participation

Table 5 Urban–rural and male–female differences in the signaling value of education

	OLS		Heckman		OLS		Heckman	
	Urban (1)	Rural (2)	Urban (3)	Rural (4)	Male (5)	Female (6)	Male (7)	Female (8)
High school	- 0.144 (0.418)	0.055 (0.101)	- 0.089 (0.405)	0.136 (0.113)	0.034 (0.136)	0.020 (0.132)	0.089 (0.147)	0.089 (0.147)
College	- 0.263 (0.792)	0.380** (0.166)	- 0.016 (0.770)	0.468** (0.185)	0.172 (0.230)	0.267 (0.210)	0.262 (0.255)	0.336 (0.239)
Bachelor's	0.104 (0.924)	0.667*** (0.182)	0.385 (0.897)	0.771*** (0.206)	0.505* (0.257)	0.678*** (0.236)	0.599** (0.289)	0.743*** (0.276)
Master's	0.107 (1.314)	1.099*** (0.340)	0.590 (1.280)	1.159*** (0.348)	0.820*** (0.394)	0.914** (0.363)	0.890** (0.434)	1.044** (0.416)
Education	0.104 (0.127)	0.029 (0.022)	0.038 (0.126)	0.055* (0.031)	0.037 (0.033)	0.044 (0.029)	0.035 (0.039)	0.034 (0.041)
Family background	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mills:lambda	-	-	- 0.610*	0.223*	-	-	0.071	- 0.035
N	467	686	471	4159	617	536	2119	2511
R ²	0.296	0.301	-	-	0.268	0.355	-	-

1. Standard errors are reported in parentheses. *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. 2. After testing, the independent variable's average VIF value is less than 10, and there is no serious multicollinearity. 3. The sample size of the Heckman model reported by Stata, different from OLS, includes those samples with missing values for the income variable. Therefore, the sample size is larger than that of OLS

there is indeed the sample selection bias according to the inverse mills ratio at the 10% level, information asymmetry in the rural is even more significant, and the above conclusion is still valid.

Besides, the signaling value of education for females is usually greater than that for males. In particular, bachelor's signaling values for females are 93.1%, respectively. By contrast, males are only 60.2% relatively. According to the inverse mills ratio, there is no significant sample selection bias in the Heckman model.

Considering that the sheepskin effect is more significant in rural and females, we compare the signaling value of different income groups in rural and females to explain further the reasons for the changes in the proportion of mechanism. Simultaneously, to strengthen the conclusion's conviction, we exclude extreme income samples (below 10% and above 90% quintile).

Table 6 shows that the signaling values of the college, bachelor's, and master's degrees in rural-born are 14.5%, 18.5%, and 1.4%, respectively, and all are significant in the low-income group at the 5% or 10% level. By contrast, none of the sheepskin effects in the high-income group is significant. The above results indicate that low-income groups face more severe information asymmetry than high-income groups in the labor market, so the proportion of post-market discrimination channels in low-income groups is greater than that in high-income groups. At the same time, the results of columns (3)–(4) also explain, to some extent, how post-market discrimination channels account for a large proportion of male–female opportunity inequality.

6.2 Employer prejudice

We illustrate the post-market discrimination manifested as employer prejudice through the inequality in job opportunities, as shown in Table 7:

Firstly, we set the dummy variable of the workplace. The value is “1” when the employee works in a state agency or state-owned enterprise, and the value is “0” when workplaces are others, such as private enterprises. The results are shown in Column (1). Secondly, we divide the type of job into white-collar and blue-collar occupations¹¹ and set dummy variables. The white-collar is “1,” and the blue-collar is “0.” The results are shown in Column (2). Thirdly, testing robustness. We subdivide the workplace into government, state-owned enterprises, private enterprises, foreign-funded enterprises, and social organizations. The results are shown in Columns (3)–(6).

Table 7 shows that the apparent work threshold caused by employer prejudice in the labor market and employer prejudice's effect on workplaces is more significant than the type of occupation. The odds of urban and male employees entering government or state-owned enterprises are about 1.8 times and 1.35 times that of rural and female employees. At the same time, the occupation type is also affected by employer

Footnote 10 continued

but almost has no effect on their wages, according to the existing research (see Mroz 1987; Wooldridge 2000; Schwiebert 2015). Hence, we consider it the identification variable.

¹¹ Combining related research, white-collar employees are: “persons in charge of communist party, government, enterprises, and institutions,” “personnel and related personnel,” and “professional and technical personnel.” Blue-collar employees are: “manufacturing and related personnel,” “social production service and life service personnel,” “production and auxiliary personnel of agriculture, forestry, animal husbandry and sideline fishery,” and “other employees who are inconvenient to classify.”

Table 6 Differences in the signaling value of education by income group

	Rural		Female	
	Low-income (1)	High-income (2)	Low-income (3)	High-income (4)
High school	0.023 (0.048)	0.072 (0.107)	0.019 (0.061)	0.441** (0.221)
College	0.158** (0.075)	0.134 (0.196)	0.159* (0.088)	0.871** (0.417)
Bachelor's	0.193** (0.086)	0.279 (0.220)	0.270*** (0.101)	1.029** (0.483)
Master's	0.207* (0.118)	0.489 (0.342)	0.100 (0.193)	1.448** (0.681)
Education	0.000 (0.009)	- 0.014 (0.030)	0.001 (0.011)	- 0.123* (0.066)
Family background	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes
<i>N</i>	295	242	215	194
<i>R</i> ²	0.156	0.128	0.159	0.166

1. Standard errors are reported in parentheses. *, **, and *** denote significance at 10%, 5%, and 1% levels, respectively. 2. After testing, the independent variable's average VIF value is less than 10, and there is no serious multicollinearity

prejudice, but the influence is not statistically significant. To some extent, the threshold caused by prejudice prevents the social mobility of rural or female employees, causing them to linger in low-income jobs for a long time, which partly explains why post-market discrimination channels account for a higher proportion.

7 Conclusions

Equality of opportunity is an essential foundation of social justice. Based on CLDS data, we measure the degree of opportunity inequality and its influencing factors. Then, we extend the Oaxaca–Blinder approach to estimate pre-market and post-market discrimination to explain the influence mechanism of IO. Besides, we present the mechanism formation from information asymmetry and employer prejudice further. The conclusions are as follows:

(1) IO is a fundamental cause of employees' income inequality, explaining 14.45% of the total value. IO for females is more significant than for males, and IO for the rural is greater than for the urban. What's more, IO is primarily manifested in urban–rural and male–female opportunity inequality, accounting for 17.16% and 31.66%, respectively. (2) From the influence mechanism, urban–rural opportunity inequality is mainly

Table 7 Urban–rural and male–female differences in job opportunities

	Logit model			Multinomial logit model		
	State-owned or government	White-collar	Government	State-owned	Private	Foreign
	(1)	(2)	(3)	(4)	(5)	(6)
Hukou	1.795*** (0.289)	0.823 (0.131)	1.189 (0.348)	1.877** (0.582)	0.741 (0.210)	0.784 (0.323)
Gender	1.353** (0.189)	1.107 (0.151)	0.763 (0.176)	1.000 (0.248)	0.531*** (0.112)	0.706 (0.211)
Family background	Yes	Yes	Yes	Yes	Yes	Yes
Personal characteristics	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	1131	1199	1131	1131	1131	1131
<i>R</i> ²	0.162	0.157	0.125	0.125	0.125	0.125

Standard errors are reported in parentheses. ** and *** denote significance at 5% and 1% levels, respectively

realized through the pre-market discrimination channel, while post-market discrimination is the primary channel of male–female opportunity inequality. By comparison, urban–rural opportunity inequality comes more from the pre-market effect of human capital accumulation. Moreover, the primary mechanism for IO has not changed by region. Although the young can get an education more equitably, they suffer severer post-market discrimination, especially the employer prejudice against gender reflected in the labor market. (3) Due to variances in information asymmetry and employer prejudice, the influence mechanism has different characteristics in different income groups. With the increase in income, the proportion of the post-market discrimination channel shows a downward trend. The possible explanation is that the more significant information asymmetry and employer prejudice will lead to employees facing invisible employment thresholds, making them stay in the lower-income jobs for a long time.

Based on the above, the policy enlightenment lies in: on the one hand, equalizing public resources and increasing educational input in backward areas to eliminate pre-market discrimination. It is indispensable to guarantee fairness in vulnerable groups' educational opportunities, especially the children in rural. On the other hand, the government should improve the dual structure of urban and rural areas and break the hukou barriers to facilitate cross-regional mobility of labor. It is necessary to ensure equal male–female employment opportunities by eliminating post-market discrimination. Besides, to solve the problem of inequity from the source, the government needs to improve the labor market by removing information asymmetry and employer prejudice.

Acknowledgements We sincerely thank the editors for their hard work, and we are also very grateful to anonymous referees for their helpful suggestions. We are especially indebted to China Labor-force Dynamic

Survey (CLDS) for providing data support. The opinions are the author's alone. Please refer to <http://css.sysu.edu.cn> for more information about the CLDS data.

Funding The funding was provided by National Social Science Foundation of China (Grant No. 18BJY213).

Declarations

Conflict of interest No conflict of interest exists in submitting this manuscript, and all authors approve the manuscript for publication. On behalf of my co-authors, I would like to declare that the work described was original research that has not been published previously and is not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

References

- Arrow KJ (1973) The theory of discrimination. In: Ashenfelter O, Rees A (eds) *Discrimination in labor markets*. Princeton University Press, Princeton, pp 3–33
- Bauer TK, Dross PJ, Haisken-DeNew JP (2005) Sheepskin effects in Japan. *Int J Manpow* 26(4):320–335. <https://doi.org/10.1108/01437720510609528>
- Björklund A, Jäntti M, Roemer JE (2012) Equality of opportunity and the distribution of long-run income in Sweden. *Soc Choice Welfare* 39(2–3):675–696
- Bourguignon F, Ferreira FH, Menéndez M (2007) Inequality of opportunity in Brazil. *Rev Income Wealth* 53(4):585–618
- Cappelen AW, Sørensen EØ, Tungodden B (2010) Responsibility for what? Fairness and individual responsibility. *Eur Econ Rev* 54(3):429–441
- Checchi D, Peragine V (2010) Inequality of opportunity in Italy. *J Econ Inequal* 8(4):429–450
- Chetty R, Hendren N, Kline P, Saez E (2014) Where is the land of opportunity? The geography of inter-generational mobility in the United States. *Q J Econ* 129(4):1553–1623
- Coate S, Loury GC (1993) Will affirmative-action policies eliminate negative stereotypes? *Am Econ Rev* 83(5):1220–1240
- Conde-Ruiz JJ, Ganuza JJ, Profeta P (2022) Statistical discrimination and committees. *Eur Econ Rev* 141:103994
- Davillas A, Jones AM (2020) Ex ante inequality of opportunity in health, decomposition and distributional analysis of biomarkers. *J Health Econ* 69:102251
- Démurger S, Gurgand M, Li S, Yue X (2009) Migrants as second-class workers in urban China? A decomposition analysis. *J Comp Econ* 37(4):610–628
- Donni PL, Peragine V, Pignataro G (2014) Ex-ante and ex-post measurement of equality of opportunity in health: a normative decomposition. *Health Econ* 23(2):182–198
- Fajardo-Gonzalez J (2016) Inequality of opportunity in Adult Health in Colombia. *J Econ Inequal* 14(4):395–416
- Ferreira FH, Gignoux J (2011) The measurement of inequality of opportunity: theory and an application to Latin America. *Rev Income Wealth* 57(4):622–657
- Ferrer AM, Riddell WC (2002) The role of credentials in the Canadian Labour Market. *Can J Econ* 35(4):879–905
- Fleurbay M, Schokkaert E (2011) Equity in health and health care. In: Pauly MV, McGuire TG, Barros PP (eds) *Handbook of health economics*. North-Holland, Waltham, pp 1003–1092
- French D (2012) Causation between health and income: a need to panic. *Empir Econ* 42(2):583–601
- Fuller S, Vosko LF (2008) Temporary Employment and social inequality in Canada: exploring intersections of gender, race and immigration status. *Soc Indic Res* 88(1):31–50
- Golley J, Zhou Y, Wang M (2019) Inequality of opportunity in China's labor earnings: the gender dimension. *Chin World Econ* 27(1):28–50
- Heckman JJ, Kautz T (2013) Fostering and measuring skills: interventions that improve character and cognition. *SSRN Electron J*

- Jaeger DA, Page ME (1996) Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *Rev Econ Stat* 78(4):733–740. <https://doi.org/10.2307/2109960>
- 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 Econ* 22(12):1470–1495
- Krafft C, Alawode H (2018) Inequality of opportunity in higher education in the Middle East and North Africa. *Int J Educ Dev* 62:234–244
- Kübler D, Schmid J, Stüber R (2018) Gender discrimination in hiring across occupations: a nationally-representative vignette study. *Labour Econ* 55:215–229
- Lang K, Manove M (2011) Education and labor market discrimination. *Am Econ Rev* 101(4):1467–1496
- Lefranc A, Pistoletti N, Trannoy A (2008) Inequality of opportunities vs. inequality of outcomes: Are western societies all alike? *Rev Income Wealth* 54(4):513–546
- Li CL (2005) Prestige stratification in the contemporary China: occupational prestige measures and socio-economic index. *Sociol Stud* 02:74-102+244 (in Chinese)
- Li Y, Lv G (2019) Research on the source and channel of opportunity inequality in China. *China Ind Econ* 9:60–78 (in Chinese)
- Liu Q (2012) Unemployment and labor force participation in urban China. *China Econ Rev* 23(1):18–33
- Marrero GA, Rodríguez JG (2013) Inequality of opportunity and growth. *J Dev Econ* 104:107–122
- Meng X, Zhang J (2001) The two-tier labor market in urban China. *J Comp Econ* 29(3):485–504
- Mora JJ, Muro J (2008) Sheepskin effects by cohorts in Colombia. *Int J Manpow* 29(2):111–121. <https://doi.org/10.1108/01437720810872686>
- Mroz TA (1987) The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica* 55(4):765–799
- Neal DA, Johnson WR (1996) The role of premarket factors in black-white wage differences. *J Polit Econ* 104(5):869–895
- Olfindo R (2018) Diploma as Signal? Estimating sheepskin effects in the Philippines. *Int J Educ Dev* 60:113–119
- Ordine P, Rose G (2011) Inefficient self-selection into education and wage inequality. *Econ Educ Rev* 30(4):582–597
- Palomino JC, Marrero GA, Rodríguez JG (2019) Channels of inequality of opportunity: the role of education and occupation in Europe. *Soc Indic Res* 143(3):1045–1074
- Piketty T (1995) Social Mobility and redistributive politics. *Q J Econ* 110(3):551–584
- Roemer JE (1993) A pragmatic theory of responsibility for the egalitarian planner. *Philos Public Aff* 22(2):146–166
- Roemer JE (1998) Equality of opportunity. Harvard University Press, Cambridge
- Rosa Dias P (2009) Inequality of opportunity in health: Evidence from a UK cohort study. *Health Econ* 18(9):1057–1074
- Schady NR (2003) Convexity and sheepskin effects in the human capital earnings function: recent evidence for Filipino men. *Oxford Bull Econ Stat* 65(2):171–196
- Schwiebert J (2015) Estimation and interpretation of a Heckman selection model with endogenous covariates. *Empir Econ* 49(2):675–703
- Sewell WH, Shah VP (1968) Parents' education and children's educational aspirations and achievements. *Am Sociol Rev* 33(2):191
- Shi X, Wei L, Fang S, Gao X (2018) Opportunity inequality in China's income distribution. *Manage World* 3:27–37 (in Chinese)
- Solon G (1999) Intergenerational mobility in the labor market. *Handbook of Labor Economics*, pp 1761–1800
- Spence M (1973) Job market signaling. *Q J Econ* 87(3):355
- Spera C, Wentzel KR, Matto HC (2008) Parental aspirations for their children's educational attainment: Relations to ethnicity, Parental Education, children's academic performance, and parental perceptions of school climate. *J Youth Adolesc* 38(8):1140–1152
- Wooldridge JM (2000) Introductory econometrics a modern approach. South-Western, Cincinnati
- Yamamoto Y, Matsumoto K, Kawata K, Kaneko S (2019) Gender-based differences in employment opportunities and wage distribution in Nepal. *J Asian Econ* 64:101131
- Yang X, Gustafsson B, Sicular T (2021) Inequality of opportunity in household income, China 2002–2018. *China Econ Rev* 69:101684
- Zhang L, Sharpe RV, Li S, Darity WA (2016) Wage differentials between urban and rural-urban migrant workers in China. *China Econ Rev* 41:222–233

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.