I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14463

Inequality of Opportunity in Household Income, China 2002-2018

Xiuna Yang
Björn Gustafsson
Terry Sicular

JUNE 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14463

Inequality of Opportunity in Household Income, China 2002-2018

Xiuna Yang
China Development Research Foundation

Björn Gustafsson
University of Gothenburg and IZA

Terry Sicular
University of Western Ontario

JUNE 2021

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Inequality of Opportunity in Household Income, China 2002-2018*

This study contributes to the literature on inequality of opportunity (IOp) in China by covering a longer and more recent span of time, employing better measures of given characteristics, and analyzing IOp for household income per capita with comparisons to individual income. Furthermore, we study how IOp differs between the rural- and urban-born, and how IOp changes across birth cohorts and with age. We use 2002, 2013 and 2018 data from the Chinese Household Income Study and focus on income inequality among working-age persons. Between the years of study, IOp in China declined, especially between 2013 and 2018. In 2002 the large contributors to IOp were region, hukou type at birth, and parents' characteristics. In 2018 the contributions of region, hukou type at birth and parents' occupation had decreased, but that of parents' education had increased. We find that IOp is larger among those born in rural than urban China. Furthermore, IOP's contribution to total inequality within each birth cohort is highest earlier in individuals' work lives and declines with age. IOp is higher for older than younger birth cohorts, reflecting that younger cohorts have benefited from increased opportunities associated with China's reforms and opening up.

JEL Classification: D31, D63, J62

Keywords: Gini, income inequality, inequality of opportunity, China

Corresponding author:
Björn Gustafsson
Department of Social Work
University of Gothenburg
P.O. Box 720
SE 405 30 Göteborg
Sweden

E-mail: Bjorn.Gustafsson@socwork.gu.se

* The authors would like to thank participants at the International Workshop of the China Household Income Project 2018, Hangzhou, November 14-15, 2020, and referees to this journal for helpful comments and suggestions. This work was supported by the Faculty of Social Science, University of Western Ontario .

1. Introduction

There are now many studies of income inequality at the household level in China. Most show that inequality was rising until around 2008, then levelled off or declined. This inequality has been associated with various underlying factors such as the rural/urban divide, education, the unequal distribution of property and assets, and government transfer and tax policies, to mention a few. In this paper we look at underlying factors through a different lens: the inequality of opportunity (IOp).

While the literature on income distribution and inequality now has a long history, the literature on IOp is young. Much of the work on this topic is inspired by Roemer's (1998) observation that some inequality is due to factors under the control of the individual because of his or her *efforts* and *choices*, and some inequality is due to factors that are beyond the individual's control, that is, *circumstances* and *opportunities*. The first reflects inequality for which people can be held responsible, and the second reflects inequality for which they should not be held responsible. This raises an empirical question: How much of observed inequality in income is due to circumstances and opportunities, and how much is due to efforts and choices? The answer to this question is of clear policy relevance. If a large share of income inequality is due to circumstances, then an argument exists for policy interventions.

The empirical literature on IOp is now extensive and includes some studies of China. The Chinese case is interesting for several reasons. China is a country that has experienced rapid economic growth. Rapid economic growth can expand opportunities and open pathways for greater economic mobility. Also, China is an interesting case because it has been characterized by substantial barriers to mobility between the rural and urban sectors. These barriers are associated with the household registration or hukou system. Over time China has adopted a series of policy reforms to reduce the extent of hukou restrictions. In principle, then, opportunities for China's rural population should have increased, causing a decline in IOp.

Our analysis makes several contributions. First, our period of study is longer and more recent than that of previous studies of IOp in China. We analyze changes in IOp between 2002, 2013 and 2018, spanning a longer and more recent time frame than previous studies of IOp for China. These years saw not only substantial economic growth and reforms reducing hukou barriers, but also structural changes following China's accession to the WTO, a major expansion of tertiary and secondary education, the Global Financial Crisis, China's large economic stimulus response to that crisis, and a substantial broadening of China's social welfare system. Also, in the latter part of this period China began a transition from rapid economic growth to a 'new normal' characterized by slower growth. Our analysis provides new information about how IOp in China has evolved in the context of these changes.

Second, we examine the heterogeneity of IOp among groups that have different circumstances. We investigate IOp for rural versus urban groups, where rural and urban are defined by hukou at birth. Some existing studies also examine IOp separately for China's rural and urban populations (e.g., Shi et al. 2018, Ma and Zou 2018, Li and Lv 2019). These studies differ from ours in that they use current hukou, not hukou at birth (an exception is Golley, Zhou and Meng 2019). Use of hukou at birth is preferred, because the analysis can then reflect opportunities to change hukou, which have increased in China over time.

We also examine IOp heterogeneity among different age groups and birthyear cohorts. This analysis sheds light on whether IOp in China changes systematically with age, and whether IOp differs between younger and older cohorts. Ours is one of the few analyses for China of IOp by birthyear cohort.

We use household income per capita as the target variable for the analysis. In this regard our approach is consistent with general studies of inequality and related policy discussions, which typically focus on household income. Most studies of IOp, however, use individual earnings as the target variable. Use of individual income makes sense in IOp analysis because the subject of interest is the individual. Also, in many countries where the majority of income is earned through wage employment, individual income is readily identified.

For reasons that we discuss more fully later, we prefer household income per capita to individual income as the target variable for our analysis. One reason is that a substantial share of China's population earns income through farming and household business, activities in which the income of individual family members is not easily identified. Another reason is that individuals' circumstances can affect their opportunities for marriage and household formation, have repercussions on intrahousehold arrangements, and have spillover effects on the earnings of other family members and of their households as a whole.

In China such circumstance effects could be important. Studies find that marriage choices in China are influenced by circumstances such as parents' education, parents' occupation, and hukou (see, for example, Zhou 2019). Furthermore, decisions about marriage, fertility and the intrahousehold division of labor are determined not only by the interests of individuals, but also by the collective interests and capabilities of the family (Zang and Zhao 2018). Such circumstance effects are not unique to China. Charles, Hurst and Killewald (2012), for example, find that in the U.S. parental wealth is more important for marital sorting than the individuals' own education.

In view of these considerations, we use household income per capita as the target variable for our base analysis and for our analyses of IOp heterogeneity. For the base analysis, however, we also present estimates using individual income. Comparison of the results for household versus individual income yields some insights into the sensitivity of estimates of inequality and IOp to the choice of target variable.

We conduct our analysis using repeated cross-section data from the 2002, 2013 and 2018 rounds of the China Household Income Project (CHIP) surveys. In order to exclude the effects on inequality of choices regarding schooling and retirement, we restrict our sample to working-age individuals. As a consequence, our findings differ somewhat from those of studies of the whole population.

Turning to results, we find that IOp in China has declined. The decline is especially marked from 2013 to 2018. Furthermore, we find that the circumstance variables that are the most important contributors to IOp have changed. In 2002 region, hukou type at birth, and parents' characteristics were the largest contributors. As of 2018 the contributions of region, hukou at birth and parents' occupation had decreased, and that of parents' education had increased.

With respect to heterogeneity of IOp, we find that IOp is larger for the rural-born than for the urban-born. Furthermore, the importance of IOp declines with age, that is, as members of a particular birthyear cohort age, the proportion of within-cohort income inequality that can be attributed to IOp decreases. We conclude that IOp is indeed heterogeneous among groups; however, such compositional differences do not explain the decline in China's overall IOp. Instead, the decline is mainly associated with declines over time in the IOp of individual circumstance variables such as hukou type at birth and region, as well as some decline in between-group IOp.

We begin in Section 2 with a selective review of relevant literature. Section 3 discusses our methods and data. Section 4 reports our base results. Here we discuss changes in total inequality and IOp for China as a whole and examine the importance of the different circumstance variables. In Section 5 we present alternative results using individual income as the target variable and discuss how choice of target variable affects the results. In Sections 6 and 7 we explore heterogeneity of IOp based on hukou type at birth, age, and birthyear cohort. In Section 8 we sum up the study and draw some lessons.

2. Relevant Literature: A Selective Overview

The literature on IOp has grown rapidly and there are now many studies that attempt to quantify the importance of IOp.[1] This growth is understandable, as the IOp approach has several attractive properties.[2] Perhaps most importantly, it allows childhood's long-run consequences to be characterized by more than one variable. Furthermore, it is flexible with respect to the measure of inequality. The approach also has some drawbacks. As applied so far, empirical analyses of IOp implicitly assume that the relationship between circumstances and the outcome

[1] For example, the survey by Roemer and Trannoy (2015), Ramos and Van de gaer (2016), and Ferreira and Peragine (2016).

[2] See Björklund and Jäntti (2020), who compare the IOp approach with the related literatures on intergenerational mobility, intergenerational effects and sibling correlations.

is causal, which might not always be the case. It should also be understood that to take full advantage of the approach, rich data is required.

Answering the question of how large a proportion of inequality is due to opportunities, and how large a proportion is due to efforts and choice, depends on several factors. One is obviously the choice of population or society that is studied, and when it is studied. Another is research design. In the empirical literature different target variables have been analyzed. Many studies have focused on earnings among individuals. Others have, as here, investigated disposable household income per capita, with individuals as the unit of analysis. Other studies have focused on alternative welfare-relevant variables such as the length of education or health.[3]

Another factor regards the structure of the methodology applied. One approach is to disaggregate the population under study on the basis of a few variables, and to decompose the value of inequality indices by cells defined by these variables. Another, and probably more frequently applied, method is, as here, to use a regression analysis framework – the parametric approach. A third variation in method regards the set of circumstance variables included in the study.

A common criticism of the IOp literature arises because the estimated importance of effort is given by the 'unexplained' variation in the target variable after controlling for circumstance variables. Consequently, studies that have inadequate data on circumstance variables will tend to understate the inequality of opportunity (Peragine 2004, Zhou and Zhao 2019). This aspect of the methodology means that having relatively complete, good data on circumstance variables is important.

Finally, there is the question of which inequality index to use when summarizing the distribution of income. For example, when the Gini coefficient is used, the results typically show a larger proportion of total inequality is due to circumstances than when using indices that belong to the entropy family such as the Theil index or the Mean Logarithmic Deviation. Researchers make different methodological choices in these regards, which means that results from different studies can be difficult to compare.

One early and much cited study that aimed to quantify the importance of IOp is Bourguignon et al. (2007). These authors analyzed how in earnings of urban persons aged 26 through 60 years in Brazil were related to race, region of birth, and parental education. The results showed that 10 to 37 percent of inequality as measured by the Theil index could be attributed to circumstances. They also showed that parent's education made larger contributions to IOp than either race or place of birth. In a similar analysis Ferreira and Gignoux (2011) broadened the analysis to cover six Latin American countries using harmonized data for persons aged 30 through 49. The authors reported that between one fourth and one half of earnings inequality in those countries was attributed to inequality of opportunities.

[3] On the latter, see, for example, Trannoy et al. (2010).

Two articles, one by Brunori, Ferreira and Peragine (2013) and another by Roemer and Trannoy (2015) provide surveys of the research on cross-country differences in IOp. They find that among high-income countries, the level of IOp is lowest in the Nordic countries and highest in Ireland, Spain, the U.K., and the U.S. The highest levels of IOp in the world, however, are in developing countries, for example, in Brazil, Guatemala, Panama and South Africa. A broad conclusion from the available cross-country evidence is that “All in all, it seems that the inequality of opportunities (across countries) is highly correlated with inequality of income” (Roemer and Trannoy 2015, p 293). Neither of these studies contains estimates for China.

With respect to estimates of IOp for China, this work is in general not covered in the published reviews we refer to above, but a range of studies are available. Here we discuss a selection of relevant ones. With respect to the English-language literature, an early study of IOp in China is by Zhang and Eriksson (2010). Their data was an unbalanced panel of 1,287 observations from the China Health and Nutrition Survey (CHNS) in nine provinces covering the period 1989 to 2006. They used a parametric regression estimation method, and their circumstance variables included parental background (education, employment, and income) and the children's gender, age, birthplace, education, and employment. Income inequality was measured over individual income using the Gini-coefficient.

The results indicated that China has had a substantial degree of IOp during 1989 and 2006, with its share of total income inequality rising over time from 46% in 1989 to 63% in 2006. The two most important circumstance variables were parental income and parents’ type of employer, while the role of parental education was minor. Income inequality increased during the period studied, as did IOp. Most of their estimations, however, use a sample that is disproportionately urban-based (88%) and composed of state-owned enterprise workers (67%). According to the WDI (World Development Indicators), China’s urban population during this time frame was considerably lower, initially only 26 to 45 percent. Their sample is further unrepresentative because in the CHNS dataset information on parents’ characteristics is only available for individuals residing in the same households as their parents. Consequently, the estimation sample is restricted to adults who live with their parents. In addition, this paper used the Gini index to measure IOp but did not correct the coefficients for biases caused by path-dependency (Foster and Shneyeroy, 2000).

A more recent study is by Zhou and Zhao (2019), which used a sample of about 4,000 observations also from the CHNS but for the more recent years 1993 to 2011. Similar to Zhang and Eriksson (2010), the outcome variable is individual income, they restricted the estimation sample to individuals living with their parents, used a parametric approach and the Gini coefficient, and did not correct for path dependency. Unlike Zhang and Eriksson, they restricted their analysis to males and used only three circumstance variables: region of residence, mother’s household registration (hukou), and father’s income. Not surprisingly, they found a lower level and share of IOp to total inequality than Zhang and Eriksson, but a similar increase through 2006.

From 2006 to 2011, however, IOp declined. As a share of total inequality, the share of IOp rose from 15 percent in 1993 to 21 percent in 2006, then fell to 9 percent in 2011.

Another recent English-language study for is by Wu (2018), which used a fairly large sample of 19,736 observations drawn from the China Family Panel Studies (CFPS) for 2010 and 2012. Their outcome variable was individual income, inequality index was the Gini coefficient, and, similar to our analysis, they used a parametric approach and the Shapley decomposition. They found that at the national level relative IOp was at least 30% in 2010 and 40% in 2012. The most important circumstance variables were gender, geographic characteristics, and parents' socioeconomic status.

Golley, Zhou and Meng (2019) examine IOp of labor earnings with a focus on gender differences using data for 2010 from the Survey of Women's Social Status in China. They reported relative IOp of 25 percent and found that gender was the most important circumstance variable, followed by father's education and occupation, and region. Age cohort also contributed to IOp, although IOp showed no clear pattern across age cohorts. Unusually, this study had access to information about hukou at birth (most other studies only had information on current hukou). So measured, hukou contributed about 10 percent of total IOp.

	sample	dataset, years	inequality index, outcome variable	methodology	main results
English language literature					
Zhang and Eriksson (2010)	Urban+rural	CHNS 1989, 1991, 1993, 1997; 2000, 2004, 2006	Gini, individual annual income	ex-ante; parametric; own decomposition method	1. Relative IOp is between 46% and 63% from 1989 to 2006. 2. Both IOp and income inequality are increasing during this period.
Golley. Zhou and Wang (2019)	Urban+rural	Survey of Women's Social Status in China 2010	Gini and MLD, individual labor earnings (time frame not specified)	ex-ante; parametric; Shapley decomposition	1. Relative IOp is 25%. 2. Gender, father's education and occupation, and region are the most important factors for IOp; age cohort and hukou at birth also contribute. IOp across age cohorts shows no clear pattern.
Zhou and Zhao (2019)	Urban+rural	CHNS 1993, 2000, 2006, 2011	Gini, individual annual income	Ex-ante, parametric; own decomposition method	The share of IOp rises from 15% in 1993 to 21% in 2006, then falls to 9% in 2011.
Wu (2018)	Urban+rural	CFPS 2010, 2012	Gini, individual annual income	Ex-ante; parametric; Shapley & Oaxaca decompositions	1 Relative IOP is at least 30% in 2010 and 40% in 2012. 2. Gender, geographic characteristics and parents' socioeconomic status are the three main factors.
Chinese language literature					
Li and Lv (2016)	Urban	CHIP 2007	Gini, individual monthly income	ex-ante; parametric; Shapley & Field decompositions	1. Relative IOp is more than one third. 2. Gender and region are relatively important factors for IOp.
Li and Lv (2018)	Urban	CHIP 2007, 2008	MLD, individual monthly income	ex-ante; parametric; 2SLS; no decomposition	1. Relative IOp is 23%. 2. The relationship between IOp and age is inverted U type. IOp shows an inverted U-shaped relationship with age.
Shi, et al. (2018)	Urban+rural	CGSS 2013	Gini, individual annual income	ex-ante; parametric; Shapley decomposition	1. Relative IOp is 36%. 2. Father's background (education and occupation), birthplace and hukou type are relatively important factors for IOp. 3. With the increase of age from 20 to 60, IOp gradually increases.
Li and Lv (2019)	Urban+rural	CGSS 2008, 2010, 2013, 2015	MLD, individual annual income	ex-ante; parametric; Shapley decomposition	1. Relative IOp is 46%, 39%, 42%, 35% in 2008, 2010, 2013, and 2015, respectively. 2. Region, hukou and gender are the main factors for IOp. 3. IOp is lower for post-1980s birthyear cohorts than for 1950s, 1960s and 1970s cohorts.
Chen and Huang (2015)	Urban+rural	CHNS 1989, 1991, 1993, 1997; 2000, 2004, 2006, 2009	Gini, Theil, Atkinson, individual annual income	ex-ante; parametric; no decomposition	Relative IOp is 54.61% on average from 1989 to 2009.

Gong, Li and Lei (2017)	Urban+rural	CGSS 2008, 2010, 2011, 2012, 2013	Theil, individual annual income	ex-post; non-parametric; no decomposition	Relative IOp is 39.26%, 35.01%, 42.96%, 38.69% for post-50s, post-60s, post-70s and post-80s birthyear cohorts, respectively.

Table 1: Studies of IOp in China

Table 1 also shows Chinese-language studies of IOp in China. The findings of Chinese-language studies differ, in part because they use different methodologies and inequality indexes, and they employ different datasets covering different years and different populations (rural, urban or the whole country). For example, Li et al. (2016) investigated relative IOp using the CHIP 2007 data for urban China and found that relative IOp was more than one third. Li and Lv (2018) using a different inequality index found that this proportion is 23 percent.

Several studies used the China General Social Survey (CGSS) data. Shi et al. (2018), Li and Lv (2019), and Gong et al. (2017) found that during the years 2008 and 2015 relative IOp in China was between 35 percent and 46 percent. Using the China Health and Nutrition Survey, however, Chen and Huang (2015) found that relative IOp was 55 percent on average from 1989 to 2009, with are no obvious trend over time.

As shown in Table 1, several papers have compared the IOp among different age groups. For example, Li et al. (2018) found that IOp showed an inverted U-shaped trend from ages 18 to 60. Shi et al. (2018) reported that IOp gradually increased by age from 20 to 60. A second group of papers have compared IOp across different birth cohorts. For example, Li and Lv (2019) found that IOp was lower for people born during the 1980s than for those born earlier. In contrast, Gong et al. (2017) found no clear pattern in how IOp varies across birth cohorts.

There is some agreement on the main factors contributing to IOp. Li and Lv (2016), Shi et al. (2018), and Li and Lv (2019) all found that region, hukou status, gender, and father's background (education and occupation) were relatively important factors contributing to income inequality of opportunity. An interesting study by Golley and Kong (2016) looked at IOp in education rather than income. They used the China Family Panel Studies (CFPS) dataset and concluded that hukou type was the single biggest contributor to IOp in education.

Compared with the existing literature on China, our paper makes three main contributions. First, we investigate changes in IOp over a longer and more recent time frame than previous studies. Our study covers periods of faster (2002-2013) and slower (2013-2018) economic growth, as well as periods with less and more education expansion, and with less and more barriers to spatial mobility. Secondly, the three waves of CHIP data that we use contain rich information, which allows us to incorporate relatively more and better circumstance variables than many other studies. Thirdly, the three waves make it possible for us to analyze how IOp within birth cohorts changes with age, and also how IOp differs across birth cohorts.

Most studies in the IOp literature estimate IOp in individual income or individual earnings. To our knowledge, all existing studies of China to date estimate IOp in individual income or earnings. Our approach differs from the prior literature in that we analyze IOp in household income per capita rather than individual income. Consequently, our findings provide a broader picture of income inequality that encompasses income from activities that involve multiple household members, e.g., farming, and reflects the indirect effects of circumstances through marriage, family formation, and the intrahousehold division of work.

3. Methodology and Data

Within the IOp framework, income is determined by two sets of factors: factors beyond the individual's control, which are denoted as "circumstances," and factors within the individual's control, which are denoted as "effort." Economic inequality due to circumstances is considered inequality of opportunity and associated with economic unfairness. Within this framework, we adopt the so-called parametric ex-ante approach (Ferreira and Gignoux 2011; Bourguignon et al. 2007). This approach assumes that the conditional distribution of effort is independent of circumstances.

Assuming that income Y is determined by a finite set of exogenous circumstances C that are beyond an individual's control and also by unobserved effort E, one can write the following function:

$$Y = f(C, E, u) \qquad (1)$$

Effort E is affected by circumstances C and other unobserved factors, so equation (1) can be rewritten as:

$$Y = f(C, E(C, v), u) \qquad (2)$$

Equation (2) reflects both the direct and indirect effects of circumstances on income. The indirect effects occur because circumstances affect effort, which in turn affects income.

Note that income here is usually interpreted as individual earnings, but it can also be interpreted more broadly as household income or household income per capita, in which case it encompasses the income of other household members and from joint household production. An individual who belongs to a household will derive utility from these other sources of household income. Moreover, individuals' circumstances pre-date and influence their marriage and household formation, so that these components of household income are indirectly functions of the individual's circumstances. For these reasons, in most of our analyses we use household income per capita (see below for further discussion).

Under the assumption of log linear additivity, a parametric implementation of the above model is:

$$lnY_i = \alpha + \beta C_i + \gamma E_i + \varepsilon_i \qquad (3)$$

where

$$E_i = \delta C_i + e_i \qquad (4)$$

The coefficients β and γ measure the relationship between circumstances and income, and effort and income. δ measure the relationship between circumstances and effort.

Combining equations (3) and (4), we obtain:

$$lnY_i = \alpha + (\beta + \gamma\delta)C_i + \gamma e_i + \varepsilon_i \qquad (5)$$

or, simplifying:

$$lnY_i = \alpha + \rho C_i + \omega_i \qquad (6)$$

Equation (6) can be estimated by OLS. $\hat{\rho}$ measures both the direct relationship between circumstances and income and the indirect relationship between circumstances and income through effort. For each individual predicted income from this regression equation can be interpreted as an estimate of income associated with circumstances. The difference between predicted and actual income is an estimate of income associated with effort, i.e., the relationship between effort and income is reflected in the estimated residuals.

Using the estimated parameters of equation (6), one can decompose inequality of income between inequality associated with effort and inequality associated with circumstances. The decomposition can be carried out in total and separately for each circumstance variable.

Let income inequality $\mathrm{I}(Y)$ be additively decomposable between the inequality of income associated with circumstances $\mathrm{I}(Y_c)$ and the inequality of income associated with effort $\mathrm{I}(Y_e)$:

$$\mathrm{I}(Y) = \mathrm{I}(Y_c) + \mathrm{I}(Y_e) = \sum_j I(Y_{C_j}) + \mathrm{I}(Y_e) \qquad (7)$$

Inequality of opportunity $\mathrm{I}(Y_c)$ is equal to the inequality of predicted incomes over all *i* individuals $\hat{Y}_i$, where $\hat{Y}_i = \exp(\hat{\alpha} + \hat{\rho}C_i)$, that is, inequality of opportunity is $\mathrm{I}\,(\hat{Y})$. The relative share of inequality of opportunity is $\mathrm{I}(\hat{Y})/\mathrm{I(Y)}$. Inequality of effort $\mathrm{I}(Y_e)$ is equal to all remaining inequality $\mathrm{I(Y)} - \mathrm{I}(\hat{Y})$, and its relative share is $1 - \mathrm{I}(\hat{Y})/\mathrm{I(Y)}$.

Ideally, inequality decomposition based on equation (7) should add up exactly. Also, it should be path independent, that is, the result should not depend on the order in which one carries out the decomposition across the different circumstance factors. The natural decompositions of many commonly used inequality indexes such as the Gini coefficient, however, are not path independent. Fortunately, the Shapley decomposition method provides a flexible approach that can be applied to different inequality indices without violating path independence. The Shapley decomposition is thus an attractive approach for identifying the overall effect of circumstances versus effort on inequality, as well as the individual contribution of each circumstance variable to inequality. We use the Shapley decomposition method for our analysis. Details of the Shapley decomposition method can be found in Shorrocks (2013) and Wan (2002, 2004).

In our analysis we use three alternative inequality indexes, the Gini coefficient, Theil index and MLD (Mean Logarithmic Deviation). As shown in the literature, we find that although the level of IOp differs among the different indexes, changes of IOp over time and comparisons among population groups are consistent. Therefore, we mainly report results for the Gini coefficient.

With respect to data, we use the rural, rural-to-urban migrant, and urban samples from the CHIP surveys for the years 2002, 2013, and 2018. The overall sample sizes in the three years are 63,911 for 2002, 57,821 for 2013, and 70,431 for 2018. We restrict the estimation sample in several ways. First, to maintain consistency across the waves of the CHIP survey, we use data only from the 14 province-level units common to all three years.[4] Note that in all estimations we apply two-level (region×urban/rural/migrant) population-based sampling weights to adjust for differences in sampling probabilities across years, regions and sectors.

In order to focus on opportunities during working age, we further restrict our estimation sample to persons aged 26 to 50. This means we do not study children, young adults and older people, and our estimates are therefore not affected by choices regarding length of schooling or retirement age. We further restrict the sample to household heads and spouses. This is because the CHIP surveys only collected information on parents' education and occupation for household heads and spouses. After applying these several restrictions, our estimation sample sizes are 16,249 for 2002, 13,342 for 2013, and 9,724 for 2018.

Our income variable is household income per capita. In this regard our approach differs from most other studies of IOp, but we prefer household income per capita because individuals benefit from household income that comes from sources other than their own person-specific earnings. In addition, an individual's circumstances can affect those other components of household income. Circumstances pre-date and have an effect on marriage, household formation, family structure, and intrahousehold choices, all of which can in turn influence household income.

[4] The CHIP 2002 rural, 2013 rural and urban, and 2018 rural and urban surveys covered 14 common provinces: Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Shandong, Henan, Hubei, Hunan, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu. The CHIP 2002 urban survey covered only 12 of these provinces (Hunan and Shandong were not included). Our analysis therefore uses data from the 14 common provinces except for the 2002 urban sample, which covers 12 of these common provinces.

For example, an individual who is born with a rural hukou will have different marriage opportunities than an individual born with an urban hukou. As a result, hukou at birth will affect how much income is earned by one's spouse. Similarly, individual circumstances such as parents' education can indirectly affect household choices regarding intrahousehold specialization, e.g., one household member may specialize in unpaid household responsibilities while the other specializes in paid wage employment. In such cases, individual earnings will not fully reflect the returns to the individual from his or her circumstances and can cause measurement error in the estimation of IOp.

An individual's circumstances can also affect income from joint production such as farming or a family business. Joint production is common in our estimation sample, which includes rural and rural-urban migrant households. Joint household production can involve complementarities and substitutions among household members, as well as economies of scale. The circumstances of one family member, e.g., her parents' education or occupations, can generate spillovers that affect the productivity of other family members.

A third reason for using household income per capita is practical. In household survey datasets like CHIP, some components of household income, e.g., income from household farming and family businesses, as well as some types of transfers, are not associated with individuals. Accurately assigning shares of this income to each family member is not straightforward and doing so requires extensive information that is often unavailable. Consequently, studies of IOp based on individual income often understate or mismeasure these components of income.

For these reasons we use household income per capita as the target variable. We acknowledge that household income per capita also has drawbacks and can lead to measurement error. The extent to which that measurement error is greater or less than measurement error for individual earnings will depend on the degree to which individual circumstances affect marriage and household choices and the extent of circumstance spillovers among family members. In the Chinese context, such factors are likely to be important. In view of these considerations, we compare our results for household income per capita (in Section 4) to alternative estimates for individual income, which we report in Section 5.

The CHIP datasets contain household income as defined by the National Bureau of Statistics. NBS income is the sum of labor earnings, net business income (from both farm and non-farm businesses), pension income, property income, and transfers received by the household net of transfers out, including both public and private transfers.[5,6] We divide this measure of household

[5] Note that we do not include imputed rents on owner-occupied housing in our income variable.

[6] Our household income variable mostly follows the NBS definition of income, but with some adjustments. The NBS revised its income definition in 2013. The most important revision was the addition of imputed rents on owner-occupied housing for urban households. This income component had not been included previously. In order to maintain consistency over time, we have subtracted imputed rents from the NBS income variable for 2013 and

income by the number of household members to obtain household income per capita. Income is measured in constant prices across the three years, that is, we deflate incomes using the urban consumer price index (CPI) for the urban and migrant samples and the rural CPI for the rural sample. In our calculations we do not adjust for spatial price differences in the cost of living within China in our base price year due to the lack of good spatial price indexes for China.

Following the literature, our circumstance variables are individual characteristics that individuals cannot choose or change, but that can directly or indirectly affect current incomes: gender, age, ethnicity, hukou type at birth, region, occupation of parents, and education of parents.[7] We note that, like most studies of IOp, the available set of observed circumstances is limited by data availability and some relevant circumstances will necessarily be omitted from the analysis. As discussed by Ferreira and Gignoux (2011) and others, adding additional circumstance variables increases estimated IOp. As a result, estimated levels of IOp including those reported here should be viewed as lower bound estimates.

Most, but not all, of our circumstance variables are defined consistently and are available in the CHIP survey data for all three years. One variable that is not defined consistently is parents' occupation. For 2013 and 2018, we have full information on eight occupational sub-groups, but for the 2002 rural sample we only know whether the parents have ever engaged in non-agricultural activity. To allow comparability across all three years and between the urban and rural samples, in our base estimations we define parents' occupation as a binary variable, peasant or non-peasant. For 2013 and 2018 we present alternative estimates using the eight parental occupational categories.

In principle, the region variable should identify the individual's location of residence during childhood. Unfortunately, this information is available only for 2018. For all other years we only have location of current residence. Using the 2018 data we carried out the estimation two ways, once using location of residence in childhood and then using location of current residence. The results were similar. IOp was slightly lower when using region of childhood residence than current region. This is consistent with the fact that some inequality associated with current region is due to choice (effort) rather than circumstances. Because the results using these two region variables are similar for 2018, we infer that current region of residence is a reasonable proxy for

---

2018. In 2013 the NBS also made a few other technical adjustments to its definition of income. To evaluate the effect of these other adjustments on our findings, we ran some alternative estimations of IOp for 2013 with and without such adjustments to the NBS income variable. This exercise indicated that these additional adjustments to the NBS definition of income may have reduced the levels of national and urban IOp (and their contributions to total inequality) by roughly 5-10%. Rural IOp estimates were not affected. We conclude that our estimates likely understate slightly the increases in IOp from 2002 to 2013 nationwide and in urban areas. For more details on the NBS income definition changes in 2013, see Sicular et al. (2020).

[7] Some other circumstance variables, specifically, parent's party membership and sibling order, are available for one or two years of the CHIP survey. We estimated alternate specifications including these variables in the years for which they are available. Including these additional variables did not substantially change the results. Therefore, we do not report results including these additional variables here.

location of residence in childhood in the earlier years 2002 and 2013, at which time barriers to geographic mobility in China were higher.

The descriptive statistics for our dependent and independent variables are shown in Table 2. The descriptive statistics are calculated using regional x urban/rural/migrant population weights so that they reflect the regional and urban/rural/migrant composition of China's population in each year.[8] The gender balance in all years leans slightly towards females. The average age in all years is about 40, reflecting that we restrict the estimation sample to ages 26 through 50. The sample is predominately Han ethnicity, with only 6-7 percent ethnic minority.

Because we are interested in circumstances at birth, our analysis differentiates between urban and rural based on hukou at birth versus current hukou. The majority of individuals in the samples were born with agricultural (rural) hukou and a minority with non-agricultural (urban) hukou. The share born with non-agricultural hukou is 34 percent in 2002 and 26 percent in 2013 and 2018.[9] The shares of the population born with non-agricultural versus agricultural hukou are different than the shares holding current non-agricultural versus agricultural hukou. This difference is evident in our estimation samples. In 2002 the share of the estimation sample holding current non-agricultural hukou was 44 percent, larger than the share born with non-agricultural hukou (34 percent). In 2013 the shares were 46 and 28 percent, and in 2018 the shares were 48 and 26 percent, respectively.[10] The higher share of current non-agricultural (urban) hukou is explained by hukou conversions and the reclassification of some rural places (and their populations) as urban.

[8] Note that these weights are based on current location of residence and current hukou. Rural means resident in rural areas with agricultural hukou. Urban means resident in urban areas with non-agricultural hukou. Migrant means resident in urban areas with agricultural hukou.

[9] The decline in the share with non-agricultural hukou at birth from 2002 to 2013 reflects the impact of China's family planning policies. Individuals in our 2002 estimation sample were born between 1952 and 1977, before the one-child policy and mostly before the 1970s, when China first embarked on birth control policies with the 'late, sparse, few' program. From the start, China's birth control policies have imposed stricter limits on the urban than the rural population. As a result, the rate of natural increase of the urban-born population has been slower than that of the rural-born population, causing a decline in the share of the population born with non-agricultural hukou.

[10] Calculated with weights. Note that we classify individuals with the new *jumin* or citizen hukou as having current urban hukou. The new citizen hukou policy was initiated after 2002 and expanded over time. Under this policy, the agricultural and non-agricultural hukou distinction is eliminated within individual jurisdictions, and people in these jurisdictions then have a uniform local citizen hukou (Song 2014). In 2013 10 percent and in 2018 19 percent of our estimation samples held citizen hukou.

Variable		2002			2013			2018		
		Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Dependent variable										
household income per capita per day (yuan)		16249	14.265	12.825	13342	53.000	52.665	9724	83.771	77.158
ln(household income per capita per day)		16249	2.347	0.797	13342	3.661	0.814	9724	4.083	0.885
individual income per day (yuan)		12215	25.914	24.637	10153	104.606	114.413	7964	163.985	168.854
ln(individual income per day)		12215	2.785	1.174	10153	4.379	0.779	7964	4.779	0.851
Independent variables (circumstances)										
gender	Male	16249	0.477	0.499	13342	0.459	0.498	9724	0.480	0.500
	Female	16249	0.523	0.499	13342	0.541	0.498	9724	0.520	0.500
age group	Age	16249	40.266	6.381	13342	41.334	6.284	9724	39.942	6.667
	age squared	16249	1662.055	506.400	13342	1747.989	499.297	9724	1639.779	517.583
ethnicity	Han	16249	0.930	0.255	13342	0.938	0.241	9724	0.940	0.238
	ethnic minority	16249	0.070	0.255	13342	0.062	0.241	9724	0.060	0.238
hukou type when born	agricultural hukou	16249	0.661	0.473	13342	0.743	0.437	9724	0.744	0.436
	non-agricultural hukou	16249	0.339	0.473	13342	0.257	0.437	9724	0.256	0.436
region	eastern region	16249	0.406	0.491	13342	0.424	0.494	9724	0.370	0.483
	central region	16249	0.334	0.472	13342	0.307	0.461	9724	0.355	0.479
	western region	16249	0.260	0.439	13342	0.269	0.444	9724	0.268	0.443
education of parents	father's education year	16249	4.469	3.839	13342	4.852	3.842	9724	5.964	3.615
	the square of father's education years	16249	34.710	47.760	13342	38.308	44.972	9724	48.636	46.268
	mother's education year	16249	2.191	3.183	13342	3.409	3.600	9724	4.468	3.715
	the square of mother's education years	16249	14.931	33.561	13342	24.586	35.132	9724	33.761	40.247
occupation of parents	both parents peasants	16249	0.581	0.493	13342	0.573	0.495	9724	0.532	0.499
	one parent not peasant	16249	0.065	0.246	13342	0.080	0.272	9724	0.067	0.251
	both parents not peasant	16249	0.354	0.478	13342	0.347	0.476	9724	0.401	0.490

Table 2: Descriptive Statistics for Variables in Regression Model (national estimation sample)

Notes:

1) Household income per capita is measured at the household level. Individual income and all the independent variables are measured at the individual level.

2) The number of observations for individual income are smaller than for household income per capita due to missing values of individual income for some observations in the sample. The missing values mainly occur among observations that are rural or that are spouses of household heads.

3) In the 2002 rural sample we only know whether the parents have been engaged in industry and commercial business or not. We define the "occupation of parents" variable in this way: If they were never engaged in industry and commercial business, their occupation is peasant; if they have been engaged in industry and commercial business, their occupation is not peasant. For the 2002 urban sample and both the rural and urban samples in 2013 and 2018 we have more information about parents' occupation and can identify whether the parents worked in agriculture or non-agriculture.

4) The estimation sample is restricted to adults aged 26 and 50 and to household heads and spouses (parents' information is only available for household heads and spouses).

5) All means and standard deviations are calculated using region x urban/rural/migrant population weights. The number of observations shown here is not weighted.

Most individuals in the samples had one or both parents who were peasants. Parental education was low but increased over time. These parental characteristics are not surprising given the age of the parents of the individuals in the samples. The youngest observations in the 2018 sample were born in 1992. China's marriage law sets minimum age at 20 for women and 22 for men, so the parents of the youngest observations would have been born at least 20 years before 1992, that is, in 1972. Consequently, the parents of individuals in our samples would have completed their education no later than the early 1990s.

We estimate the income equations using OLS with region times urban/rural/migrant weights that reflect the regional and urban/rural/migrant composition of China's population in each year. Because the estimation sample contains individuals residing in the same households, we adjust the standard errors for clustering by household.

4. Base Results: Nationwide IOp

Table 3 shows our base estimates of total inequality, inequality of opportunity, and the ratio of inequality of opportunity to total inequality (IOR) for 2002, 2013 and 2018.[11] We report estimates for three alternative indexes: the Gini coefficient, the Theil index, and the Mean Log Deviation (MLD). The levels of IOp are different for the three indexes, but all show a decline over time in both the level of IOp and the IOR. In later sections we also carried out all analyses using the three indexes, and the main findings for the Theil and MLD are again similar to those for the Gini. For this reason, in later sections we report results only for the Gini coefficient. This facilitates comparisons with other studies of income inequality in China, most of which use the Gini coefficient.

The IOR is highest for the Gini and lowest for the Theil index, with the MLD in between. In 2002, for example, IOp contributes 40% of the Gini, 38% of the MLD, and 25% of the Theil. This is a fairly wide range and reveals the sensitivity of IOp estimates to the choice of inequality index, a reflection of the different weights that each index places on incomes in different segments of the income distribution.

Despite these differences, estimates for all three indexes point to a clear, common conclusion: between 2002 and 2018 IOp in China declined both in absolute terms and as a share of overall inequality. For the Gini the level of IOp is stable from 2002 to 2013 and then declines by 28% from 2013 to 2018. The IOR of the Gini declines from 40% in 2002 to 37% in 2013 and then further to 26% in 2018. For the MLD the level of IOp declines from 2002 to 2013 and again from 2013 to 2018; the IOR drops from 38% in 2002 to 24% in 2013 and further to 18% in 2018.

---

[11] Results of the nationwide and also the separate rural and urban income regressions are reported in Appendix Table A-1.

	2002	2013	2018
Gini:			
A.Total Inequality	0.392	0.426	0.440
B. Inequality of Opportunity	0.156	0.156	0.113
C. IOR: Ratio of Inequality of Opportunity to Total Inequality (B/A)	39.9%	36.7%	25.7%
Theil:			
A.Total Inequality	0.261	0.328	0.337
B. Inequality of Opportunity	0.064	0.067	0.042
C. IOR: Ratio of Inequality of Opportunity to Total Inequality (B/A)	24.6%	20.5%	12.5%
Mean Log Deviation (MLD):			
A.Total Inequality	0.311	0.309	0.345
B. Inequality of Opportunity	0.119	0.075	0.063
C. IOR: Ratio of Inequality of Opportunity to Total Inequality (B/A)	38.2%	24.3%	18.1%

Table 3: Nationwide Total Inequality and Inequality of Opportunity

Notes: Calculated using the Shapley decomposition method. Regression results are shown in the Appendix Table A-1.

The shrinking contribution of IOp to overall inequality is partly due to declines in the level of IOp and partly due to rising total inequality. Total inequality as measured by the Gini coefficient increases by 9 percent from 2002 to 2013 and another 3 percent from 2013 to 2018, for an overall increase of 12 percent. Similarly, from 2002 to 2018 the Theil and MLD indexes increase by 29% and 14%, respectively.

Trends in our estimates of total inequality differ somewhat from trends in estimates published by the NBS. NBS estimates of the Gini coefficient show an increase in nationwide inequality from 2002 through about 2008, after which inequality declined slightly but with no clear trend between 2013 and 2018. Unlike the NBS estimates, our estimates of the Gini for 2013 and 2018 indicate an increase in national inequality of 4.8 percent.

The increase in our estimates between 2013 and 2018 can be entirely explained by our choice of income variable and by the restricted age range of our estimation sample. If we include all ages and use the original NBS income variable in the CHIP dataset, then the Gini coefficients for 2013 and 2018 are almost the same. When we replace the original NBS income variable with our income variable, which excludes the imputed rent component of NBS income, then the Gini in 2018 is 2.4 percent higher than in 2013. If in addition we restrict the estimation sample to ages 26 through 50, then the Gini in 2018 is 4.8 percent higher than in 2013.

What is the contribution of the different circumstance variables to overall inequality? Estimates of IOp decomposed by circumstance variable are reported in Table 4. For brevity, here and hereafter we report estimates only for the Gini coefficient. The results for the Theil and MLD show similar trends over time.

	2002	2013	2018
Gender	0.1	1.0	0.3
Age	1.3	2.0	2.5
Ethnicity	0.5	0.9	0.4
Hukou type when born	9.3	7.6	3.3
Region of residence	9.3	10.1	4.7
Parents' years of education	7.6	9.8	9.7
Occupation of parents	11.8	5.4	4.8
IOR: Ratio of Inequality of Opportunity to Total Inequality	39.9	36.7	25.7

Table 4: Contributions of Circumstance Variables to Nationwide Inequality, 2002-2018 (% of total inequality)

Note: The results shown here are for the Shapley decomposition of the Gini coefficient. As discussed in the text, due to lack of data on place of residence in childhood for 2002 and 2013, for those two years our estimates are for current region of residence. The 2018 dataset contains information both on current region and region at age 14, and the results using the two variables were similar. For 2018 the estimates are for region of residence at age 14. This applies to the estimates reported in all tables and figures.

First, we note that gender and ethnicity contribute 1% or less of total inequality in all three years. We also note that the estimated regression coefficients for these two variables are not always significant. The small contribution of gender is not surprising given that we measure inequality of household income per capita, which removes differences in income between male and female members of the same household. The contribution of age is also small and not always significant, although increasing over time. We look in more detail at the relationship between age and IOp below. We note again that our analysis is restricted to the working-age population and so does not capture inequality associated with children or the elderly.

The literature on China has placed much emphasis on the importance of hukou to one's life chances. Individuals born with non-agricultural hukou have considerable advantages compared to those born with agricultural hukou (see, for example, Gustafsson, Yang and Sicular 2020). Hukou is consistently significant in our regressions. Our IOp estimates indicate that in 2002 hukou at birth contributed nearly 9% of total inequality, a substantial share. The importance of hukou at birth declined over time, however, and especially between 2013 and 2018. By 2018 hukou at birth contributed only 3% of total inequality. This decline likely reflects ongoing reductions in hukou restrictions and increased geographical mobility (Chan and Wei 2019, Zhang, Wang and Lu 2019). Again, we note that our analysis is for the working-age population only. The importance of hukou for IOp would likely be larger if retirees were included because of the large difference in pension incomes between those with non-agricultural and agricultural hukou.

Most studies of IOp for China use current hukou rather than hukou at birth as the circumstance variable because most datasets do not contain information about hukou at birth. The CHIP datasets contain information on both current hukou and hukou at birth. We therefore carried out an alternative estimation of IOp using current hukou to see how this affects the results. We found that IOp estimated using current hukou gave a higher total level of IOp and also increased the IOR. This difference in results reflects the fact that some rural-born individuals were able to obtain urban hukou.

Region of residence is also significant and contributes about 10% of total inequality in both 2002 and 2013, but its contribution drops to 5% in 2018. This decline reflects increased geographic mobility as well as ongoing reductions in inequality between regions (Luo, Sicular and Li 2020; Luo, Li, Yue and Sicular 2020).

Our results indicate that parents' characteristics are among the most important circumstance variables during the years of study. The estimated coefficients on parents' characteristics are significant for all years for the full sample, but not always significant for the separate urban and rural subsamples. In 2002 parents' education contributes 8% and in 2013-2018 about 10% of total inequality in the full sample. In 2002 parents' occupation contributes 12% and in 2013-2018 about 5% of total inequality in the full sample. Together, the parental characteristics contribute more than 20% of total inequality in 2002, and about 15% in 2013 and 2018 in the full sample.

These estimates highlight the importance of intergenerational transmission of opportunities. Two words of qualification are, however, needed. First, in our sample parents' characteristics and hukou at birth are correlated. We unfortunately do not have information about parents' hukou registration, but the parents of people born with non-agricultural hukou were likely to themselves have held a non-agricultural hukou. Consequently, the parents' education levels and occupations are to some degree a function of their hukou. Our estimates of IOp for parents' characteristics therefore reflect the influence of the parents' hukou.

Second, these estimates are based on regressions in which parents' occupation is measured by a dummy variable that only indicates whether the parents' main occupation is agriculture or non-agriculture. We use this simple indicator of parents' occupation because the 2002 data, as discussed above, do not contain any other information about the occupation of rural parents. This indicator of occupation does not capture the increasing importance and multiplicity of non-agricultural occupations in both rural and urban China in recent years.

Fortunately, the 2013 and 2018 CHIP surveys collected richer information about parents' occupation, and for those years we are able to conduct the analysis using multiple non-agricultural occupation categories. Table 5 reports the results of this alternative analysis side-by-side with our basic results from Table 4.

Not surprisingly, using a finer, more disaggregated occupation variable increases the contribution of parents' occupation and of overall IOp to total inequality. The increases, however, are small, especially in 2013. In 2013 the contribution of parents' occupation increases by less than one percentage point, and the IOR increases trivially by only 0.1 percentage point. In 2018 the increases are larger but still small.

We conclude that for China as a whole, using the basic occupation variable does not substantially bias the results in earlier years, but may cause a small understatement of the importance of IOp in 2018. Unless noted otherwise, the estimates reported in the remainder of the paper use the basic version of the occupation variable.

	2002, Basic	2013, basic	2013, alternative	2018, basic	2018, alternative
Gender	0.1	1.0	1.0	0.3	0.2
Age	1.3	2.0	2.0	2.5	2.2
Ethnicity	0.5	0.9	0.8	0.4	0.4
Hukou type when born	9.3	7.6	7.3	3.3	4.4
Region of residence	9.3	10.1	10.1	4.7	5.1
Parents' years of education	7.6	9.8	9.3	9.7	10.0
Occupation of parents	11.8	5.4	6.3	4.8	6.6
IOR: Ratio of Inequality of Opportunity to Total Inequality	39.9	36.7	36.8	25.7	28.8

Table 5: Contributions of Circumstance Variables to Nationwide Inequality using Alternative Parents' Occupation Variables, 2002-2018 (% of total inequality)

Note: The results shown here are for the Shapley decomposition of the Gini coefficient. The variable for parents' occupation in the basic results distinguishes between two occupation categories, agriculture and non-agriculture. The alternative variable for parents' occupation distinguishes among eight occupation categories: 1) peasants; 2) The responsible person of State organizations, Party organizations, enterprises, institutions; 3) Professional and technical personnel 4) Clerical/office staff, and related personnel; 5) Business / commerce, service personnel; 6) Production, transport equipment operators and related personnel; 7) soldiers; 8) others. The results shown here are for the decomposition of the Gini coefficient.

5. Alternative Results: Using Individual Income as the Target Variable

Table 6 reports alternative estimates of IOp using individual income instead of household income per capita as the target variable. We measure individual income as the sum of all income that accrues to the individual plus a share of joint household income, the latter being set equal to the per capita amount of joint household income.[12]

Comparisons between Tables 3 and 6 show that the levels of inequality for individual income are different than for household income per capita, but not predictably so. In 2002 total inequality of individual income is a bit larger, and in the other years lower, than inequality of household income per capita. The relationship between inequality for these income variables is not predictable because inequality depends on both the mean and dispersion of income. Individual income has a higher mean as well as larger dispersion than household income per capita (Table 2).

IOp of individual income is higher than that of household income per capita in 2002, lower in 2013, and about the same in 2018. Similarly, IOR is higher for individual income than household income per capita in 2002, lower in 2013, and about the same in 2018.

Despite these differences in levels, the trends across the years are similar: IOp declines for both individual income and household income per capita. The IOR also declines for both. The IOR of individual income falls from 33 percent in 2002 to 27 percent in 2018; for household income per capita it declines from 40 percent to 26 percent.

	2002	2013	2018
A.Total Inequality	0.409	0.383	0.412
B. Inequality of Opportunity	0.136	0.120	0.112
C. IOR: Ratio of Inequality of Opportunity to Total Inequality (B/A)	33.2%	31.3%	27.1%

Table 6: Nationwide Inequality of Opportunity versus Total Inequality of Individual Income

Notes: Results reported here are for the Gini coefficient.

Estimates of the contributions of individual circumstance variables to total inequality of individual income are reported in Table 7. For most circumstance variables the contributions are reasonably close to those for household income per capita (Table 4). The largest difference is for gender. For household income per capita gender's contribution to inequality of household income per capita is very low at 1 percent or less. For individual income its contribution is substantial and increases markedly over time, from 3 percent in 2002 to 10 percent in 2013 and 2018. In the latter two years gender is the single largest contributor to total inequality of individual income.

[12] In other words, the total amount of joint household income from all sources is divided by the number of household members.

The importance of gender for inequality of opportunity in individual income is not surprising. As discussed earlier, men and women in the same household have identical household income per capita. For individual income, men and women in the same household will typically have unequal income. Intrahousehold inequality will be especially large if men and women within households specialize between unpaid and paid work.

	2002	2013	2018
Gender	3.2	9.9	10.0
Age	0.6	1.7	3.2
Ethnicity	0.0	0.4	0.4
Hukou type when born	5.1	4.8	2.5
Region of residence	8.4	5.8	1.3
Parents' years of education	6.5	5.7	6.6
Occupation of parents	9.5	3.0	3.1
IOR: Ratio of Inequality of Opportunity to Total Inequality	33.2	31.3	27.1

Table 7: Contributions of Circumstance Variables to Nationwide Inequality of Individual Income, 2002-2018 (% of total inequality)

Note: The results shown here are for the Shapley decomposition of the Gini coefficient.

The choice between these two income variables for IOp analysis is a matter of judgment and will depend upon the context. Use of household income per capita is consistent with the view that the indirect returns to one's circumstances through marriage, household formation, and household decisions regarding fertility, children's schooling and labor force participation, co-residence with parents, and intrahousehold labor choices are important. Use of individual income is consistent with the view that these indirect returns are not so important, and that individual income adequately captures the returns to an individual's circumstances. The truth lies somewhere in between.

Our results for China indicate that the choice between these two income variables has some effect on the estimated levels of inequality, but both yield the same major conclusion: IOp and the IOR decline. The main difference emerges in the decomposition, specifically, in the contribution of gender. This should be kept in mind when interpreting our other findings.

6. Heterogeneity between Agricultural and Non-Agricultural Hukou at Birth

Our basic results indicated that in China as a whole, hukou at birth contributes substantially to inequality of opportunity. How important is inequality of opportunity within those born with agricultural (hereafter 'rural-born') versus those born with non-agricultural hukou (hereafter ‘urban-born’)? And, which variables are the most important contributors to IOp within each of these groups?

Because we classify individuals based on hukou at birth, in our analysis the rural-born population includes those individuals who were born with agricultural hukou and later were able to migrate and pursue opportunities in cities, including some individuals who converted to non-agricultural hukou. Including such individuals gives a fuller picture of inequality of opportunity for the rural born population than estimates based on current hukou.

As shown in Table 8, for the urban-born population in the age range of study the Gini of total inequality has risen over time but remains low. The level of IOp is also low, less than 0.10 in all three years. As a share of total urban born inequality, the contribution of IOp or IOR declines noticeably from about 30% in 2002 to 20% in 2018.

Within the urban born population the only circumstance variables with noticeable levels of IOp are region and parents’ education. Region of residence contributed more than 10% of total urban inequality in 2002 and 2013, but only 6% in 2018. The decline in the importance of region of reflects increased geographic mobility as well as economic catch up of cities in less-developed regions.

	level of IOp/inequality			% of total urban inequality		
	2002	2013	2018	2002	2013	2018
Gender	0.002	0.001	0.001	0.7	0.2	0.2
Age	0.011	0.002	0.004	3.4	0.7	1.2
Ethnicity	0.000	0.001	0.000	0.1	0.2	0.1
Region of residence	0.048	0.046	0.021	15.0	12.7	5.7
Parents' years of education	0.030	0.034	0.045	9.4	9.4	12.3
Occupation of parents	0.002	0.007	0.001	0.5	1.9	0.2
Inequality of Opportunity	0.093	0.090	0.072	29.1	25.0	19.7
Total Urban Inequality	0.321	0.359	0.363	100.0	100.0	100.0

Table 8: Urban Born Inequality of Opportunity, 2002-2018

Note: Urban born is defined as having non-agricultural hukou at birth. The results shown here are for the Shapley decomposition of the Gini coefficient.

The level and the share of IOp associated with parents’ education increase from 2002 to 2018. In 2002 and 2013 parents’ education contributes 9% of total urban-born inequality; in 2018, it contributes 12%. By 2018, parents’ education had become the most important circumstance variable for urban-born IOp. This finding underscores the importance of the intergenerational transmission of education in urban China.

IOp for all other circumstance variables is small or trivial. The low level of IOp for parents' occupation shown in Table 8, however, is partly due to the fact that these calculations use the basic version of the occupation variable that only differentiates between agricultural and non-agricultural occupations. Below we report estimates using the alternative occupation variable that differentiates among multiple non-agricultural occupations (Table 10).

Total inequality for China's rural-born population is higher than for the urban-born population (Table 9). In this respect our findings are consistent with other studies that are based on current hukou (Shi et al. 2018, Ma and Zou 2018, Li and Lv 2019). With respect to changes between the years, our estimates show that total rural-born inequality, like total urban-born inequality, increased from 2002 to 2018. Over the same period the level of rural-born IOp first increased a bit and then fell back to its initial level. Consequently, the IOR initially rose from 25% in 2002 to 29% in 2013, and then fell to 22% in 2018.

	level of IOp/inequality			% of total rural inequality		
	2002	2013	2018	2002	2013	2018
Gender	0.002	0.007	0.003	0.4	1.7	0.6
Age	0.003	0.014	0.014	0.7	3.3	3.1
Ethnicity	0.004	0.005	0.003	1.0	1.3	0.6
Region of residence	0.039	0.041	0.024	9.7	10.0	5.2
Parents' years of education	0.019	0.035	0.034	4.7	8.5	7.5
Occupation of parents	0.036	0.019	0.023	9.0	4.7	5.1
Inequality of Opportunity	0.102	0.121	0.100	25.6	29.4	22.1
Total Rural Inequality	0.403	0.411	0.451	100.0	100.0	100.0

Table 9: Rural Born Inequality of Opportunity, 2002-2018

Note: Rural is defined as having rural (agricultural) hukou at birth. The results shown here are for the Shapley decomposition of the Gini coefficient.

With respect to the different circumstance variables, region, parents' education, and parents' occupation had the largest levels of IOp. Region contributed 10% of total rural inequality in both 2002 and 2013, and 5% in 2018. As is the case for the urban-born population, the decline in rural-born IOp associated with region of residence reflects increased geographic mobility plus the economic catch up of less-developed regions.

IOp associated with parents' education increased from 2002 to 2013 and then remained basically unchanged to 2018. As a share of total rural-born inequality, IOp of parents' education rose from 5% in 2002 to about 8-9% in 2013-2018. Meanwhile, the IOp of parent's occupation followed the opposite pattern, first decreasing and then increasing slightly. In 2002 parents' occupation contributed 9% of rural born inequality, falling to 5% in 2018.

As of 2018, parents' education was the most important circumstance variable for the rural born, followed closely by region and parents' occupation. As in urban born China, then, for the rural-born population the intergenerational transmission of education has remained important for life chances.

It is possible that the estimated contributions of parents' occupation reported in Tables 8 and 9 could be understated due to the coarse occupation variable. We therefore present estimates in Table 10 based on the alternative parents' occupation indicator that distinguishes among eight occupational categories. The alternative indicator makes little difference for the rural-born IOp estimates, but it has a noticeable effect on the urban-born estimates, especially in 2018.

For the urban-born population, using the alternative occupation indicator noticeably increases the contribution of parents' occupation. In 2018, for example, using the alternative occupation indicator increases the contribution of parents' occupation to total urban-born inequality from 0.2% to 4.8%. This leads to an increase in the urban-born IOR from roughly 20% to 25%. IOp for the other circumstance variables does not change much.

	Urban				Rural			
	2013, basic	2013, alt.	2018, basic	2018, alt.	2013, basic	2013, alt.	2018, basic	2018, alt.
Gender	0.2	0.2	0.2	0.2	1.7	1.6	0.6	0.5
Age	0.7	0.7	1.2	1.3	3.3	3.2	3.1	2.6
Ethnicity	0.2	0.2	0.1	0.0	1.3	1.2	0.6	0.6
Region of residence	12.7	12.6	5.7	6.2	10.0	10.0	5.2	5.5
Parents' years of education	9.4	7.8	12.3	12.0	8.5	8.2	7.5	7.6
Occupation of parents	1.9	4.4	0.2	4.8	4.7	5.3	5.1	6.3
Inequality of Opportunity	25.0	25.9	19.7	24.5	29.4	29.5	22.1	22.9

Table 10: Estimates of Urban Born and Rural Born IOp using Alternative Parents' Occupation Variable, 2002-2018 (% of total urban born or total rural born inequality)

Note: Urban/rural is defined by hukou at birth. The alternative variable for parents' occupation used here distinguishes among eight occupation categories (see Table 5). The results shown are for the decomposition of the Gini coefficient.

How do the separate estimates for the urban and rural born relate to our nationwide estimates? Nationwide IOp was higher than IOp in either subgroup. This was due to the existence of inequality of opportunities between the rural-born and urban-born subgroups. In addition, the decline in the nationwide level of IOp between 2002 and 2018 was larger than the decline for either subgroup. For the rural subgroup, in fact, IOp did not decline at all.

Why would nationwide IOp decline more than IOp of the subgroups? The explanation is that differences in opportunities between the rural- and urban-born subgroups narrowed. This can be seen in the estimates of nationwide IOp associated with hukou at birth (Table 4). Between 2002 and 2018 the IOp associated with hukou at birth fell substantially. IOp associated with region and parents' occupation—circumstance variables that are correlated with hukou at birth—also declined. The declines in nationwide IOp associated with these circumstance variables are evidence of the narrowing of inequality of opportunity between the rural- and urban-born.[13]

7. Heterogeneity by Age and Birth Cohort

Income inequality among individuals in the same birth cohort has been found to increase with age (Deaton and Paxson 1994). An explanation for this is the accumulated effect of idiosyncratic shocks, also known as 'luck'. With respect to IOp and age, the relationship is not obvious. If unequal opportunities due to circumstances at birth have a cumulative or compounding effect over time, then IOp could increase with age. If, however, as people grow older their outcomes increasingly reflect choices, effort and luck, then IOp could decline with age.

We investigate how income inequality and IOp change with age using the CHIP data. To our knowledge, this one of the few analyses of the heterogeneity of IOp by age for China. We begin by subdividing the sample from each wave of the CHIP data into 5-year age groups (ages 26-30, 31-35, 36-40, 41-45, 56-50). We use 5-year age groups as opposed to single-year age groups so that each age group contains a sufficient number of observations to support the estimation of within-group inequality and IOp.

For each 5-year age group in each wave of the CHIP data, we estimate the income regression. Then, for each 5-year age group we calculate income inequality and IOp.[14] IOp for each age group is equal to inequality measured over the predicted incomes of observations in that age group based on the regression estimates for that age group in that year.

For each age group in each year of the CHIP data, we identify the birthyears (1988-92, 1983-87, and so on). We can then follow each birthyear cohort group across the waves of the CHIP data. So doing, we can trace the paths of inequality and IOp for each birthyear cohort group as it ages from 2002 to 2013 and 2018. This allows us to investigate whether in China inequality and IOp change systematically with age. It also allows us to examine whether inequality for particular age groups has changed over time, e.g., whether inequality among the young has increased or decreased between 2002 and 2018.

[13] Another possible explanation is that the population share of the subgroup with lower IOp (the urban-born subgroup) increased. Our estimates, however, show no increase in the share of the urban-born population (Table 2) between 2002 and 2013. In other words, changes in the urban-rural composition of the population do not explain the decline in nationwide IOp.

[14] We also carried out the analysis using the parameters from an income regression estimated over the entire sample. The results were similar.

Figure 1 shows the results for total inequality (as measured by the Gini coefficient). In the figure, each curve represents total inequality for a different birthyear cohort group. The age of the cohort group increases as one moves from left to right. For example, the green curve shows total inequality among individuals born in 1973-1977. The Gini coefficient was 0.39 when this group was 26-30 years old (estimated using the 2002 data), 0.42 at ages 36-40 (estimated using the 2013 data), and 0.42 at ages 41-45 (estimated using the 2018 data).

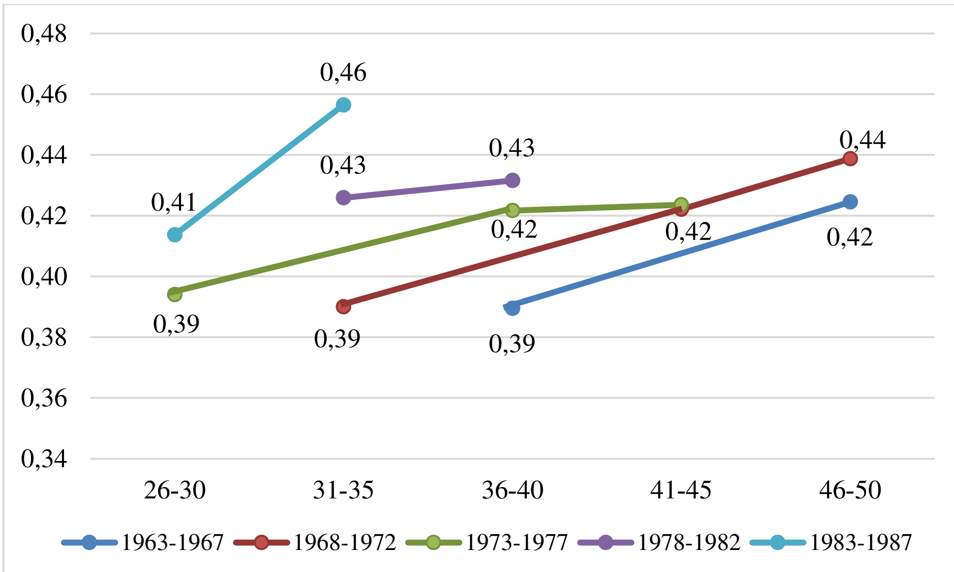

Figure 1: Income Inequality in China (Gini) by Age

The positive slopes of the inequality curves in Figure 1 indicate that within-cohort inequality increased with age. We conclude that the relationship between income inequality and age in China is similar to that observed in other countries.

Figure 1 also reveals that at any given age, within-cohort inequality is lower for older cohort groups and higher for younger cohort groups. This can be seen by the fact that the inequality curves of older cohort groups are below those of younger cohort groups. For example, at ages 36-40 inequality within the 1963-67 cohort is 0.39, within the 1973-77 cohort is 0.42, and within the 1978-82 cohort is 0.43. Put differently, inequality for younger cohorts is higher than it was for their elders when they were the same age.

Figure 2 shows a similar plot for IOp. Again, each curve represents IOp for a different cohort group, and each cohort group ages as one moves from left to right. These IOp curves are mostly downward sloping, indicating that IOp tends to decline with age. This contrasts with the positive relationship between total inequality and age.

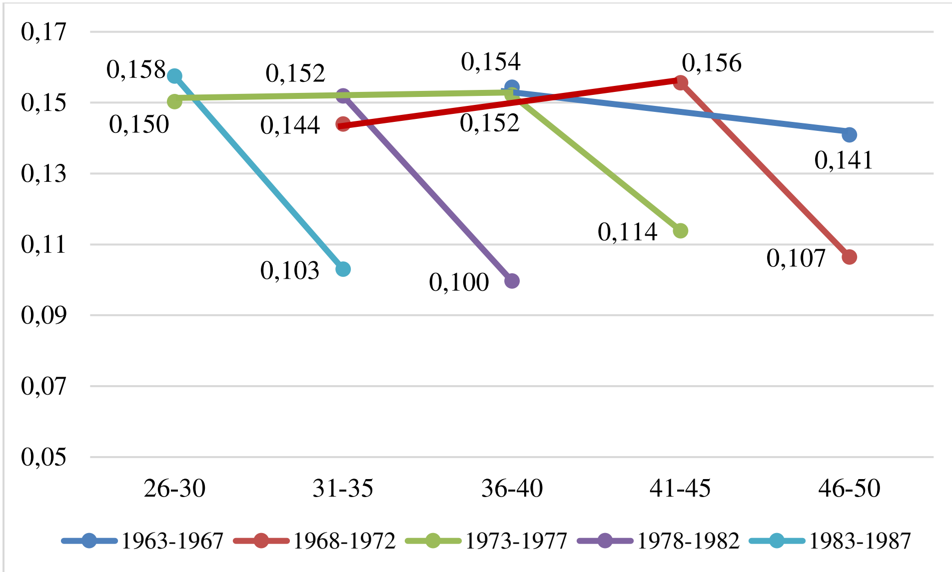

Figure 2: The Level of IOp in China (Gini) by Age

Figure 3 shows the IOR, again with each curve representing a different birthyear cohort group. All of the curves in this figure are either flat or downward sloping. Thus, IOp's contribution to total inequality within each birth cohort group is highest earlier in individuals' work lives. As individuals grow older, IOp's contribution to within-cohort inequality declines. This pattern has been observed elsewhere. As people age and progress in their careers, the cumulative effects of choices, effort and luck on outcomes grows, and the proportionate contribution of circumstances declines.

Although for each cohort group IOp and IOR tends to decline with age, IOp is higher for older cohorts than for younger cohorts. This is visible in the figures: the IOp and IOR curves of younger cohorts are to the left and below those of older cohorts. Consequently, at any particular age, the level of IOp (and IOR) is usually higher for older cohorts. For example, at ages 36-40 the IOR is 39.6 percent for the 1963-67 birthyear cohort, declines to 36.1 percent for the 1973-77 birthyear cohort, and drops even more to 23.1 percent for the 1978-82 cohort.

The higher level of IOp and IOR for older versus younger age cohorts is consistent with the decline over time in the IOp and IOR for the total population (Table 3). We speculate that this pattern reflects the policy periods in which the different cohorts grew up. The oldest cohort (born in the mid-1960s) was educated during the Cultural Revolution and early reform period, during which time opportunities were highly circumscribed and depended on one's place of birth and parents' characteristics. Over time, China's reforms and opening up policies have increased opportunities for interregional and intergenerational mobility. This is reflected in the lower IOp and IOR for younger cohorts.

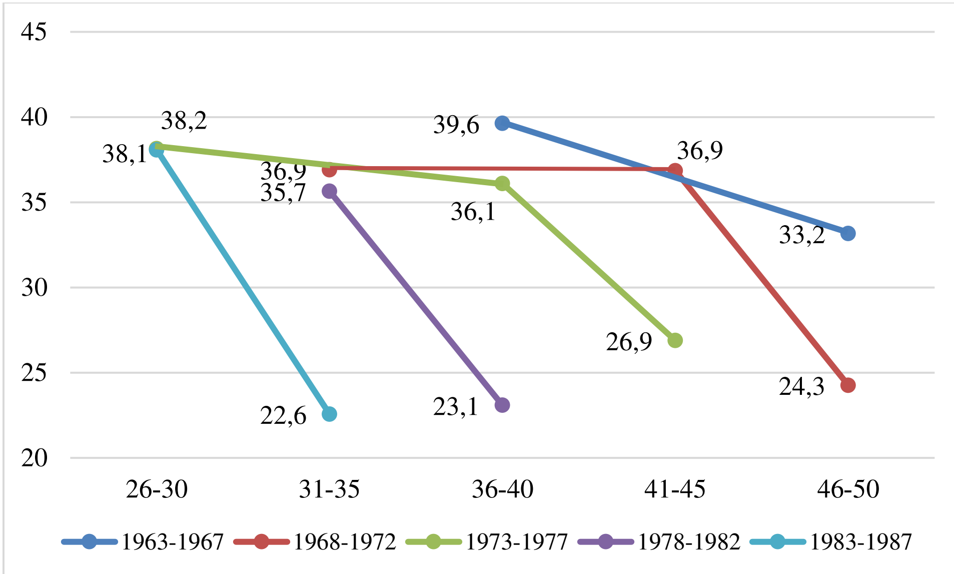

Figure 3: The Ratio of IOp to Inequality (IOR) in China, by Age (%)

Note: The results are for the Gini coefficient.

8. Conclusions

In this paper we analyze the inequality of opportunities (IOp) in China using nationwide household survey data for 2002, 2018 and 2018, years that spanned a period of dynamic change. We present estimates of IOp and the IOR (the ratio of IOp to total inequality) for China as a whole. In addition, we explore the heterogeneity of IOp between the urban-born and rural-born populations and among different birthyear cohorts.

Unlike many studies of IOp, our target variable is household income per capita. We explain that IOp analysis of household income per capita is meaningful because individuals' circumstances affect not only their individual incomes but can also have spillovers that affect the intrahousehold division of labor, individual earnings of family members, and household income from joint production. Individual income does not fully reflect these sorts of indirect circumstance spillovers. Nevertheless, use of individual income has some advantages, so we present alternative base estimates for individual income. The main difference in the results is that the contribution of gender to IOp, which is small for household income per capita, is large and increasing for individual income. The choice of target variable does not, however, alter our main findings.

Our main findings are that the level of IOp and its contribution to total inequality in China declined. Both the level of IOp and the IOR were highest in 2002 and lowest in 2018. The IOR, for example, fell from 40 percent in 2002 to 26 percent in 2018. Compared to other countries, as of 2018 the level of IOp in China was in the middle-low range, that is, lower than for most middle- and lower-income countries, somewhat higher than in the UK and US, and considerably higher than in the Nordic countries.[15]

Macroeconomic growth is usually associated with the expansion of opportunities, and slowing growth with the opposite. China's GDP grew very rapidly from 2002 to 2013, after which growth continued but at a slower pace. What was the effect of this macro slowdown on IOp? We find that the level of IOp of household income per capita was constant from 2002 to 2013 and declined from 2013 to 2018. For individual income the level of IOp declined in both periods, and more rapidly in the latter period. This pattern indicates that, contrary to expectations, the decline in China's IOp continued and perhaps accelerated during the recent period of slower macroeconomic growth.

What explains the decline in China's IOp? We obtain some answers by looking at the changing structure of IOp among characteristics and among population subgroups. Our findings suggest that much of the recent decline in IOp is associated with reductions in the IOp of specific circumstance variables. The decline also reflects that IOp has fallen across generations, that is, IOp is higher for older birth cohorts and is lower for more recent birth cohorts. We speculate that this pattern reflects that older cohorts grew up in policy periods when opportunities were limited, and that China's reforms and opening up policies have increased the opportunities for mobility for younger cohorts.

With respect to specific characteristics, the IOp associated with region of residence and hukou at birth declined. This underlines the importance of ongoing reductions in barriers to mobility, for example, through reforms of the hukou system. IOp associated with parents' occupation also declined somewhat, but IOp associated with parents' education increased. These results highlight the importance of educational legacies through the intergenerational transmission of education, as also found by Golley and Kong (2018) in their study of IOp in labor income.

The ongoing importance of parental education is somewhat surprising given the substantial expansion of both secondary and tertiary education that has taken place in China in recent decades. One reason for this may be that our sample is restricted to ages that had mostly completed their education before this expansion. Studies have found, however, that although years of education have increased for younger generations, educational persistence between parents and children associated with the rural-urban divide is ongoing (Knight et al. 2013, Rozelle and Hell 2020). In the future, inequality of opportunity associated with education will likely continue unless efforts are made to address systemic, multi-generational barriers to education, especially those faced by the rural-born population.

[15] These international comparisons are with estimates of IOp and IOR for other countries reported in Brunori et al. (2013).

References

Björklund, A. and M. Jäntti, 2020, "Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches," *Research in Social Stratification and Mobility*, Vol. 70, p. 100455.

Bourguignon, F., F. H. G. Ferreira and M. Menéndez, 2007, "Inequality of opportunities in Brazil," *Review of Income and Wealth*, Vol. 53, No. 4, pp. 585-618.

Brunori, P., F. H. G. Ferreira, and V. Peragine, 2013, "Inequality of opportunity, income inequality, and economic mobility: Some international comparisons," in E. Paus, ed, *Getting Development Right: Structural Transformation, Inclusion, and Sustainability in the Post-Crisis Era*, New York: Palgrave Macmillan US, pp. 85-115.

Chan, K. W. and Y. Wei, 2019, "Two systems in one country: The origin, functions, and mechanisms of the rural-urban dual system in China," *Eurasian Geography and Economics*, Vol. 60, No. 4, pp. 422–54.

Charles, K. K., E. Hurst and A. Killewald, 2013, "Marital sorting and parental wealth," *Demography*, Vol. 50, pp. 51-70.

Chen, D. and X. Huang, 2015, "To what extent does inequality of opportunity affect income inequality? Based on the perspective of intergenerational mobility," *Economic Review*, No. 1, pp. 3-16. (In Chinese.)

Deaton, A. and C. Paxson, 1994, "Intertemporal choice and inequality," *Journal of Political Economy*, Vol. 102, No. 3, pp. 437-467.

Ferreira, F. H. G. and J. Gignoux, 2011, "The measurement of inequality of opportunity: Theory and an application to Latin America," *Review of Income and Wealth*, Vol. 57, No. 4, pp. 622-657.

Ferreira, F. H. G. and V. Peragine, 2016, "Individual responsibility and equality of opportunity," in M. Adler and M. Fleurbaey, eds, *Handbook of Well Being and Public Policy*, Oxford: Oxford University Press, pp.746-784.

Foster, J. E. and A. A. Shneyerov, 2000, "Path independent inequality measures," *Journal of Economic Theory*, Vol. 91, No. 2, pp. 199-222.

Golley, J. and S. T. Kong, 2018, "Inequality of opportunity in China's educational outcomes," *China Economic Review*, No. 51, pp. 116-128.

Golley, J., Y. Zhou and M. Wang, 2019, "Inequality of opportunity in China's labor earnings: The gender dimension," *China & World Economy*, Vol. 27, No. 1, pp. 28-50.

Gong, F., Z. Li and X. Lei, 2017, "The effect of effort on opportunity inequality: Measurement and comparison," *Economic Research Journal,* No. 3, pp. 76-90. (In Chinese.)

Gustafsson, B., X. Yang and T. Sicular, 2020, "Catching up with the West: Chinese pathways to the global middle class," *The China Journal*, Vol. 84, pp. 102-127.

Knight, J., T. Sicular and X. Yue, 2013, "Educational inequality in China: The intergenerational dimension," in S. Li, H. Sato and T. Sicular, eds, *Rising Inequality in China: Challenges to a Harmonious Society*, New York: Cambridge University Press, pp.142-196.

Lefranc, A., N. Pistolesi and A. Trannoy, 2008, "Inequality of opportunities vs. inequality of outcomes: Are western societies all alike?" *Review of Income and Wealth*, Vol. 54, No. 4, pp. 513-546.

Li, Y. and G. Lv, 2016, "To what extent does the inequality of opportunity lead to the income inequality in urban China?" *Statistical Research,* Vol. 33, No. 8, pp. 63-72. (In Chinese.)

Li, Y. and G. Lv, 2018, "What leads to the opportunity inequality of income in urban China," *Statistical Research,* Vol. 35, No. 9, pp. 67-78. (In Chinese.)

Li, Y. and G. Lv, 2019, "Research on the source and channel of opportunity inequality in China," *China Industrial Economics,* No. 9, pp. 60-78. (In Chinese.)

Luo, C., T. Sicular and S. Li, 2020, "Overview: Incomes and inequality in China, 2007-2013," in T. Sicular, S. Li, H. Sato, and X. Yue, eds, *Changing Trends in China's Inequality: Evidence, Analysis and Prospects*, New York: Oxford University Press, pp. 35-74.

Luo, C., S. Li, X. Yue and T. Sicular, 2020, "Has inequality in China crossed the turning point? An analysis of trends in China's national inequality from 2013 to 2018," presented at China Household Income Project workshop, 9 July 2020, Beijing Normal University.

Ma, Z. and W. Zou, 2018, "Measurement and decomposition of opportunity inequality in China: Based on Counterfactual income distribution method," *Inquiry into Economic Issues*, No.11, pp. 1-9. (In Chinese.)

Morduch, J. and T. Sicular, 2002, "Rethinking inequality decomposition, with evidence from rural China," *Economic Journal*, Vol. 112, pp. 93-106.

Peragine, V., 2004, "Measuring and implementing equality of opportunity for income," *Social Choice and Welfare*, Vol. 22, No. 1, pp. 187-210.

Ramos, X. and D. Van de gear, 2016, “Approaches to inequality of opportunities: Principles, measures and evidence,” *Journal of Economic Surveys*, Vol. 30, No. 5, pp. 855-883.

Roemer, J. R. and A. Trannoy, 2015, “Equality of opportunity,” in A.B. Atkinson and F. Bourguignon, eds, *Handbook of Income Distribution*, Volume 2A, Oxford: Elsevier, pp. 217-300.

Rozelle, S. and N. Hell, 2020, *Invisible China: How the Urban-Rural Divide Threatens China’s Rise*, Chicago: University of Chicago Press.

Shi, X., L. Wei, S. Fang and X. Gao, 2018, “Opportunity inequality in China's income distribution,” *Management World,* No. 3, pp. 27-37. (In Chinese.)

Shorrocks, A. F., 2013, “Decomposition procedures for distributional analysis: A unified framework based on the Shapley value,” *The Journal of Economic Inequality*, Vol. 11, pp. 99-126.

Sicular, T., S. Li, X. Yue and H. Sato, 2020, “Changing trends in China’s inequality: Key issues and main findings,” in T. Sicular, S. Li, H. Sato, and X. Yue, eds, *Changing Trends in China’s Inequality: Evidence, Analysis and Prospects*, New York: Oxford University Press, pp. 1-34.

Song, Y., 2014, “What should economists know about the current Chinese hukou system?” *China Economic Review*, Vol. 29, pp. 200-212.

Trannoy, A., S. Tubeuf, F. Jusot and M. Devaux, 2010, “Inequality of opportunities in health in France: A first pass,” *Health Economics*, Vol. 19, No. 8, pp. 921-38.

Wan, G., 2004, “Accounting for income inequality in rural China: A regression-based approach,” *Journal of Comparative Economics*, Vol. 32, pp. 348-363

Wu, D., 2018, *Inequality of Opportunity in China*, doctoral thesis submitted for the Doctor of Philosophy Degree at the University of Queensland, https://core.ac.uk/download/pdf/189933279.pdf, accessed December 12, 2020.

Zang, X. and L. X. Zhao, 2017, *Handbook on the Family and Marriage in China*, Cheltenham, UK: Edward Elgar Publishing.

Zhang, Y. and T. Eriksson, 2010, “Inequality of opportunity and income inequality in nine Chinese provinces,” *China Economic Review*, Vol. 21, No. 4, pp. 607-616.

Zhang, J., R. Wang and X. Lu, 2019 “A quantitative analysis of hukou reform in Chinese cities 2000-2016,” *Growth and Change*, Vol. 50, pp. 201-211

Zhou, J. and W. Zhao, 2019, “Contributions of education to inequality of opportunity in income: A counterfactual estimation with data from China,” *Research in Social Stratification and Mobility*, Vol. 59, pp. 60-70.

Zhou, Y., 2019, “Economic Resources, Cultural Matching, and the Rural–Urban Boundary in China's Marriage Market,” *Journal of Marriage and Family*, Vol. 81, Issue 3, pp. 567-583.

Appendix Table A-1: Results from the Basic Nationwide, Rural and Urban Regressions

		Dependent variable: ln(household income per capita)								
		2002			2013			2018		
		national sample	rural sample	urban sample	national sample	rural sample	urban sample	national sample	rural sample	urban sample
gender	male: omitted									
	female	-0.002	-0.024***	0.037***	-0.073***	-0.090***	-0.017	-0.015	-0.024**	0.009
		(0.005)	(0.007)	(0.007)	(0.008)	(0.011)	(0.015)	(0.009)	(0.012)	(0.019)
age groups	age	-0.038***	-0.023	-0.046**	-0.059***	-0.069***	-0.010	-0.016	-0.019	0.020
		(0.014)	(0.017)	(0.019)	(0.018)	(0.021)	(0.032)	(0.020)	(0.023)	(0.036)
	age squared	0.001***	0.000	0.001***	0.001***	0.001***	0.000	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ethnicity	Han: omitted									
	ethnic minority	-0.179***	-0.201***	-0.020	-0.187***	-0.193***	-0.125**	-0.108***	-0.150***	0.038
		(0.030)	(0.035)	(0.041)	(0.037)	(0.044)	(0.062)	(0.041)	(0.047)	(0.073)
hukou type when born	Agricultural hukou: omitted									
	Non-agricultural hukou	0.314***			0.320***			0.226***		
		(0.022)			(0.026)			(0.028)		
region	eastern region: omitted									
	central region	-0.399***	-0.401***	-0.404***	-0.369***	-0.376***	-0.338***	-0.224***	-0.227***	-0.233***
		(0.017)	(0.022)	(0.022)	(0.022)	(0.026)	(0.037)	(0.025)	(0.029)	(0.046)
	western region	-0.378***	-0.412***	-0.290***	-0.339***	-0.369***	-0.196***	-0.291***	-0.345***	-0.114**
		(0.019)	(0.026)	(0.022)	(0.028)	(0.033)	(0.042)	(0.028)	(0.033)	(0.045)
education of parents	father's education years	0.018***	0.018***	0.019***	0.022***	0.020**	0.005	0.018**	0.014	0.006
		(0.004)	(0.007)	(0.005)	(0.007)	(0.009)	(0.012)	(0.008)	(0.010)	(0.016)
	square of father's education years	0.000	0.000	0.000	-0.000	0.000	0.000	0.000	0.001	0.001
		(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

	mother's education years	-0.011**	-0.029***	0.009	0.018**	0.016	0.010	0.024***	0.029***	-0.008
		(0.005)	(0.009)	(0.006)	(0.007)	(0.010)	(0.012)	(0.007)	(0.009)	(0.014)
	square of mother's education years	0.002***	0.004***	0.000	0.000	0.001	0.001	0.000	0.000	0.003***
		(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
occupation of parents	both parents peasants: omitted									
	One parent not peasant	0.314***	0.325***	-0.159***	0.083***	0.090***	-0.052	0.073**	0.031	0.105
		(0.026)	(0.029)	(0.054)	(0.029)	(0.032)	(0.067)	(0.036)	(0.041)	(0.081)
	Both parents not peasant	0.467***	0.629***	-0.027	0.217***	0.240***	0.110**	0.244***	0.289***	0.101
		(0.023)	(0.028)	(0.040)	(0.026)	(0.030)	(0.046)	(0.026)	(0.028)	(0.062)
constant		2.864***	2.669***	3.564***	5.013***	5.266***	4.229***	4.265***	4.351***	3.835***
		(0.271)	(0.341)	(0.390)	(0.359)	(0.416)	(0.621)	(0.390)	(0.451)	(0.712)
No. of obs		16249	9829	6420	13342	10442	2900	9724	7593	2131
adj. R^2		0.363	0.175	0.158	0.217	0.157	0.097	0.153	0.112	0.090
F statistic		415.301	130.678	60.538	130.261	63.704	17.304	96.057	52.732	13.003

Notes:

1) Decomposition estimates reported in this paper are based on the same regressions as those reported here but using centered categorical variables. The only difference between regression results using centered and non-centered categorical variables is the estimated constant term.

2) The estimation sample is restricted to individuals ages 26-50, who were not students in school, and for whom parents' information can be found in the data. Parents' information was only collected for household heads and spouses.

3) Standard errors are in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

4) We note that some additional circumstance variables are available for one or two of the three CHIP survey years, and to check robustness we carried out estimations for those years using the additional circumstance variables, specifically, for 2013 we included sibling order and for 2002 we included party membership of parents. The results of the alternative estimations for 2013 and 2002 were similar to those for the basic model. Results for the alternative estimations are available on request.

5) The regressions were estimated using region x urban/rural/migrant population weights.

6) For 2002 and 2013 the CHIP data does not contain information on region of origin. For these two years, the region variable is the location where the individual was living at the time of the survey. For 2018 information is available for both region of residence at age 14 and current

region of residence. We estimated the regression for 2018 using both; the results were similar. Here and elsewhere, all estimates for 2018 use location of residence at age 14 as the region variable, and all estimates for 2002 and 2013 use location of current residence as the region variable.