Working Paper Series

Inequality of opportunity and household education expenditures: Evidence from panel data in China

Yang Song
Guangsu Zhou

ECINEQ WP 2017 - 439

ECINEQ 2017 - 439
May 2017

www.ecineq.org

Inequality of opportunity and household education expenditures: Evidence from panel data in China

Yang Song[†]
Renmin University of China

Guangsu Zhou
Nankai University, China

Abstract

This paper This paper offers the first empirical evidence on the impact of inequality of opportunity on household education investment by using the by using the by using the by using the by using the by using the panel data from China Family Panel Studies (CFPS) in three in three in three in three waves (2010, 2012 and 2014). Our result suggests that inequality of opportunity has a negative effect on household education expenditures. This result is robust to robustness checks. Furthermore, the disadvantaged households (whose householders with less education, income, and rural hukou status) seem to be affected more by inequality of opportunity within the county they live in. Higher inequality of opportunity in the comparison group may reduce their incentives to investment more on education. Policy suggestions to reduce inequality of opportunity may include reducing labor market discrimination based on gender and hukou, balancing education resources to create more equal educational opportunities, and offering children education subsidies in low-income families.

Keywords: income inequality, inequality of opportunity, education expenditures, economic growth, China.

JEL Classification: J24, D33, O15, O53.

[†]**Contact details**: ys337@ruc.edu.cn (Y. Song, corresponding author), zhouguangsu@nankai.edu.cn (G. Zhou).

1. Introduction

In recent years, the concept of inequality of opportunity and its measurement have attracted much attention in economics literature. Dworkin (1981), Arneson (1989), and Sen (1985) are among a number of influential authors who have argued that inequality of opportunity, rather than inequality of outcome (such as income) should be used as the appropriate criterion for assessing the fairness of a given allocation or social system. In their opinions, if the inequality of some outcomes is generated by differences in individuals' efforts, then this type of inequality may be good for the society since it can motivate people to work harder. In contrast, if the inequality is due to factors beyond people's control such as family background, this type of inequality should be ethically unacceptable.

Roemer (1998) incorporates the concepts above into an economics model and divides the factors determining income into two categories: those people can control (called "efforts", e.g. how hard one works), and those beyond people's control (called "circumstances", e.g. race, gender, or family background). Given this distinction, he defines "inequality of opportunity" essentially as the extent to which important outcomes—such as income—are determined by circumstances beyond people's control (Ferreira and Gignous, 2011). According to this distinction, economists have developed a set of methods to empirically measure inequality of opportunity in different countries, such as ex-ante versus ex-post measures, parametric versus non-parametric estimations, and etc (Ferreira and Gignoux, 2011； Marrero and Rodriguez, 2012; Bourguignon et al., 2013).

Such a distinction between inequality of opportunity and the more standard concept of inequality of outcome is not only relevant from a normative point of view, but also has important implications. For example, a growing amount of empirical evidence shows that preferences for redistribution are shaped by fairness concerns which may in turn be influenced by the degree of inequality of opportunity (Ramos and Van de gaer, 2015). For instance, Alesina and La Ferrara (2005) show that for the United States those people who believe that individual economic success is related to individual effort rather than family background have lower preferences for redistribution.

Compared to inequality of a particular outcome (such as income), a more interesting appeal of the concept of inequality of opportunity is that it may have different effects on aggregate economic outcomes, such as economic growth. In its World Development Report of 2006, the World Bank argues that income inequality due to circumstances may lower growth, while income inequality due to responsibility-related variables may be helpful for growth (World Bank, 2005). Marrero and Rodriguez (2013) have empirically verified the prediction above using data for the U.S. from the Panel Survey on Income Dynamics in 1970, 1980 and 1990. They find that income inequality due to effort enhances income growth, while the part of income inequality which is accounted for by circumstances correlates negatively with growth.

The present paper asks a deeper question of why inequality of opportunity can lower the economic growth. The most important mechanism claimed by the existing literature is through human capital accumulation and education investment. That is, inequality of opportunity in income may lead to less accumulation of human capital and thus to lower growth, while income inequality due to efforts may encourage individuals to invest more in human capital and exert the largest effort possible (World

Bank, 2005; Marrero and Rodriguez, 2013).

Although the importance of human capital accumulation as an engine of economic growth and development has been widely recognized in theoretical and empirical studies (Lucas, 1988; Mankiw et al., 1992; Barro, 2001), the evidence on the effect of inequality of opportunity on human capital accumulation is still very rare. To our knowledge, only one theoretical attempt has been made to examine this potential relationship. Mejía and St-Pierre (2008) build a theoretical model which suggests that a higher degree of inequality of opportunity is associated with lower average level of human capital and a lower fraction of individuals investing in human capital.

Our present paper contributes to the literature by offering the first empirical evidence on the impact of inequality of opportunity on household education investment using a nationwide panel dataset from China. We select China as the research setting to conduct this study for two reasons. First, as is well known, income inequality in China has risen sharply in the past two decades, and the nation-wide Gini coefficient of individuals' annual income has increased from 0.37 in 1997 to 0.47 in 2014 (Song, 2017). Income inequality and its consequences on economy have attracted much attention in China (Meng et al., 2005; Benjamin et al., 2011; Song, 2013). Second, the degree of inequality of opportunity in China has been found to be much higher than that in OECD countries and comparable to Latin America countries where total income inequality is higher than China (Zhang and Eriksson, 2010; Song, 2017).

Our paper is related to at least two strands of literature. First, our paper adds to the literature on whether and why income inequality may hurt economic growth. Although the question of whether inequality may have a negative effect on economic growth has been asked many times, there is no consistent conclusion in the literature (Ferreira et al., 2014). Theory provides ambiguous predictions: whereas higher inequality may lead to faster growth through some channels (such as higher aggregate savings when a greater share of income accrues to the rich), it may have detrimental effects through other channels (such as lower aggregate rates of investment in human capital). Several studies predict that inequality of opportunity is bad for growth (Galor and Zeira, 1993; Krueger, 2012), and one empirical attempt has verified this theoretical argument using the US data (Marrero and Rodriguez, 2013). However, we do not have much empirical evidence on why inequality of opportunity affects economic growth. Our paper tries to empirically investigate whether inequality of opportunity has negative impacts on growth through its effects on household education investment.

Second, our paper is related to the controversial results on how income inequality affects household education investment. One the one hand, some studies find a positive relationship between income inequality and education expenditures. Jin et al. (2011) find that income inequality measured by the provincial Gini coefficient has a positive effect on household educational expenditure in urban China. Sun and Wang (2013) adopt the measure of village-level income inequality to obtain similar results. They find that the household education expenditures is positively related to the income inequality of the village the household lived in. Their explanation is the so-called status seeking hypothesis. That is, as income inequality rises, people may save more and invest more in education in order to strengthen their ability to seek high social status in the future. Increase in income inequality makes

entering a high-status club more attractive because differences in resources between the high- and low-status groups widen (Corneo and Jeanne, 1999; Jin et al., 2011).

On the other hand, income inequality may reduce education expenditures and human capital accumulation in the context of credit constraints. The main idea is the following: relatively poor individuals don't have the means to finance the accumulation of human capital, and, because they are credit constrained, they end up either not investing in human capital or investing very little. Some theoretical studies have shown that in the presence of credit market imperfections, inequality has a long-lasting negative impact on investment in human capital (Galor and Zeira, 1993; Ferreira, 2001). In a multiple steady state framework with borrowing constraints, initial heterogeneity in wealth and certain circumstances would reduce the opportunity of accessing credit to achieve higher levels of education.

Our paper contributes to this strand of literature by disentangling the effect of inequality of opportunity on education expenditures from the effect of total income inequality on education investment. Our results suggest that even after controlling for total income inequality, inequality of opportunity itself has a negative effect on household education expenditures. In summary, this paper will utilize a nationally representative tracking survey (China Family Panel Studies) in three waves (2010, 2012 and 2014) to examine the effect of inequality of opportunity on household education expenditures. The household fixed effect model shows that the effect is significantly negative, which is robust to a series of robustness checks. Furthermore, the disadvantaged households (whose householders with less education, income, and rural hukou status) seem to be affected more by inequality of opportunity within the county they live in. Higher inequality of opportunity in the comparison group may reduce their incentives to investment more on education.

The paper proceeds as follows. Section 2 reviews the literature on inequality of opportunity. Section 3 describes the dataset, introduces the measures for inequality of opportunity used this paper, and presents the results in China. Empirical models and results on examining the effect of inequality of opportunity on household education expenditures are shown in Section 4. Section 5 provides several robustness checks to our main results, and Section 6 concludes.

2. Past relevant literature on inequality of opportunity

Many studies have attempted to make empirical use of the concept of inequality of opportunity and proposed several measures. Both non-parametric and parametric methods have been used to estimate inequality of opportunity indexes using data from different countries. The nonparametric approach, suggested by Checchi and Peragine (2010), relies on alternative partitioning of the total population and standard between-group inequality decompositions. If groups are defined by circumstance characteristics—so that they correspond to Roemer's types—then the between-group inequality can be interpreted as an "ex-ante" measure of inequality of opportunity. Conversely, if groups are defined in terms of their relative position in the effort distributions across types, then inequality within groups corresponds to an "ex-post" measure of inequality of opportunity (Ferreira and Gignoux, 2011; Hassine, 2012). Using the ex-post non-parametric estimations, Checchi and Peragine (2010) find that inequality of opportunity accounts for about 20% of overall income inequality in

Italy. However, this non-parametric method requires large data sets for accuracy (Hassine, 2012). The greater the set of circumstances, the higher the number of cells in the partition and the higher the number of cells with zero or few observations.

The parametric analysis follows the work of Bourguignon et al. (2007) and Ferreira and Gignoux (2011) to estimate opportunity inequality as the difference between observed income inequality and the inequality that would prevail if there were no differences in circumstances. Specifically, they use the inequality index for the predicted income from a regression equation with circumstances characteristics as independent variables as a level measure of inequality of opportunity denoted by IOL, and refer to its ratio to total observed inequality as the percentage measure denoted by IOR. Although the regression equation cannot include all of the circumstances variables due to the data limitation, Ferreira and Gignoux (2011) prove that IOL and IOR can safely be interpreted as lower bound estimates of overall inequality of opportunity—i.e., inequality due to all predetermined circumstances, not only to those that are observed. They show that IOR ranges from 23 percent in Colombia to 34 percent in Guatemala among Latin America countries. In contrast, the OECD countries have relatively lower level of inequality of opportunity which ranges from 10-20 percent (Lefranc et al., 2008). Using the EU-SILC database, Marrero and Rodriguez (2012) estimate and compare the inequality of opportunity of 23 European countries in 2005. It turns out that the level of inequality of opportunity is very low in these countries with the lowest 3 percent in Finland and the highest 15 percent in Ireland. Studies on measuring inequality of opportunity in China are still very rare. A new paper by Song (2017) utilizes the most recent CGSS data (China General Social Survey) and finds that 27% of total income inequality can be attributed to inequality of opportunity. This ratio is comparable to that found in most Latin America countries with high income inequality (Ferreira and Gignoux, 2011).

As is seen, the existing literature overwhelmingly estimated the level of inequality of opportunity in different countries, but rarely examined its economic consequences empirically. To our knowledge, Marrero and Rodriguez (2013) is the only published paper that investigates whether inequality of opportunity can affect economic growth. Their findings suggest that this component of inequality was negatively associated with economic growth in the United States in the 1970-2000 period. However, the underlying mechanism through which the inequality of opportunity affects growth remains unclear.

The present paper is the first attempt to explore the effects of inequality of opportunity on people's economic behavior at the micro level. Specifically, we estimate the impact of inequality of opportunity on household education investment using a nationwide panel dataset from China and attempt to identify one plausible mechanism behind the relationship between inequality of opportunity and growth using the micro evidence.

3. Inequality of Opportunity in China

This section explains the data sources used in this paper and the estimations of inequality of opportunity in China.

3.1 Data sources

This paper explores the relationship between inequality of opportunity and education expenditures by using the micro data from China Family Panel Studies (CFPS). CFPS is a tracking survey conducted by the Institute of Social Science Survey at Peking University every two years. CFPS designs questionnaires on three different levels of aggregation: communities, households and individuals, in order to keep track of China's economic development and social change.

CFPS investigates a national representative sample of households in 2010 for the first time, and it represents 95% of the total population in 25 provinces. The household sample contains 14,798 households in 635 villages/communities of 162 counties. CFPS conducts follow-up surveys in 2012 and 2014, which accounts for approximately 80% of the total sample in 2010. The household questionnaire asks a set of detailed questions about income and expenditures, which can be used to calculate comparable income and expenditures for different waves. CFPS also contains information of each adult's parents, which allows us to construct the index for inequality of opportunity. The final sample used in our paper includes 9,274 households in all three waves.

3.2 Estimation procedure of inequality of opportunity

Inequality of opportunity is estimated as the between-type (*ex-ante*) inequality component following the parametric procedure of Ferreira and Gignoux (2011), Marrero and Rodriguez (2012), and Song (2017), which allows for the inclusion of a larger set of circumstances in the database. Specifically, following the convention of the literature, we divide the determinants of individual income (denoted by w) into two categories, including circumstances (denoted by C) and efforts (denoted by E). Since circumstances are economically exogenous by definition—in the sense that they cannot be affected by individual decisions—and given that efforts may be, and generally are, influenced by circumstances, we can write the following equation.

$$w=f[C,E(C,v),u] \quad (1).$$

u and v represent other stochastic factors affecting income, such as luck (Lefranc et al., 2009). For the purpose of measuring inequality of opportunity—rather than of estimating any causal relationship between circumstances, efforts, and income—we can simply estimate a log-linearized version of the reduced form equation by OLS：

$$\ln w = C\varphi + \varepsilon \quad (2).$$

We follow three steps to construct the index for inequality of opportunity. First, we estimate equation (2) and obtain the predicted income denoted as $\hat{w}$。Given the variables in the CFPS datasets, we include gender, hukou status at 3 years old, paternal and maternal education as circumstances variables in the regression equation. Second, given that the Theil index (mean log deviation) is additively decomposable, we follow the literature convention and calculate the Theil index for the predicted income denoted by T($\hat{w}$) in order to estimate the extent to which the total

income inequality can be attributed to inequality of opportunity (Shorrocks, 1984; Bourguignon et al., 2007; Ferreira and Gignoux, 2011). Third, we calculate the index for inequality of opportunity (denoted by IO) as the ratio of the Theil index for predicted income to that for the actual income.

$$IO = \frac{T(\hat{w})}{T(w)} \quad (3).$$

We will use the measure above throughout the paper to investigate the effect of inequality of opportunity on household education expenditures. Two points need to be mentioned on this index. First, we calculate this index at the county level using the CFPS dataset. This is the first study, to our knowledge, that calculates inequality of opportunity at the county level which may have larger effect on household behavior within a closely knit social comparison group (Sun and Wang, 2013). Compared with the existing studies related to inequality of opportunity in other countries (Marrero and Rodriguez, 2013), the use of county as the aggregation level in our study creates more variations than the country or state level used by others.

Second, due to the data limitation, the regression equation cannot include all of the circumstances variables, which may make the measure for inequality of opportunity imprecise (Kanbur and Wagstaff, 2015). Ferreira and Gignoux (2011) have proved that this measure can safely be interpreted as lower bound estimates of overall inequality of opportunity. Since the focus of our paper is to investigate the effect of inequality of opportunity on household education expenditures, we can avoid entering the discussion on the preciseness of this well-used measure as long as the measure is consistent for each county in each data wave. That is, our main purpose is to calculate a consistent measure for inequality of opportunity and then use the measure to study its effect on household education expenditures. We thus include commonly-used circumstances variables in the literature, such as gender, hukou status at 3 years old, paternal and maternal education (Zhang and Eriksson, 2010; Marrero and Rodriguez, 2013; Song, 2017).

3.3 Summary statistics

Table 1 presents the summary statistics of the key variables used in this paper for all three waves. From this table, we can find that both of total inequality and inequality of opportunity reach a peak in 2012. The percentage of total income inequality that is attributed to inequality of opportunity increases from 17% in 2010 to 24% in 2012, and decreases slightly to 22% in 2014. Figure 1 visualizes these trends and shows the peaks in 2012. In addition, both the natural logarithm value of education expenditures and the ratio of education expenses to total expenditure decrease in this time period. For example, the ratio of education expenditures to the total expenditure falls from 10% in 2010 to 7% in 2014, in part because the CFPS is a tracking survey for the same households and fewer children of school age are left. This assumption is verified by the fact that the ratio of children aged below 16 has fallen in this period. Another reason for the drop is that household total expenditure has increased over these years, which further reduces the ratio of education expenditures.

Table 1 and Figure 1 Inserted Here

China has a unique household registration system called hukou in Chinese. The current hukou system requires that each Chinese citizen be assigned either rural or urban hukou in a given location. People inherit at birth the hukou status from their parents. Furthermore, one's hukou status would remain unchanged no matter where the individual physically moved, unless he or she goes through a formal procedure of hukou conversion (Song, 2014). However, the present hukou conversion policy in big cities–the destination of the majority of migrants–is almost totally geared towards the super-rich and the highly educated and is beyond the reach for the majority of rural migrants (Song, 2014). As Table 1 shows, most of the householders in our sample hold rural hukou (around 70%) including some rural-to-urban migrants who work and live in cities but still keep their original rural hukou status.

We present more information on inequality of opportunity by estimating the total inequality (measured by Theil index) and inequality of opportunity (measured by the ratio of predicted Theil index to original Theil index) in different regions. Figure 2 shows the evolution of inequality and inequality of opportunity by region: east, middle and west. The first one is for year 2010, the second one is for year 2012, and the last one is for year 2014. Interestingly, in the most recent wave (2014), the East (the richer part of China) has the lowest total income inequality, but the highest inequality of opportunity, while West (the poorer part of China) exhibits the opposite pattern. This indicates that inequality of opportunity would not automatically decrease as economic growth takes place.

Figure 2 Inserted Here

Finally, we take a glance of the relationships between total income inequality, inequality of opportunity, and education expenditures using the scatter plots. Figure 3A describes the relationship between income inequality and education expenditures at county level. Income inequality is measured by the Theil index, and education expenditures is the mean of household education expenditures at the county level. Similarly, Figure 3B describes the relationship between inequality of opportunity and education expenditures at the county level. Each point in these figures corresponds to a county. Both of the two graphs exhibit significantly negative relationships.

Figure 3 Inserted Here

4. Empirical models and results

4.1 Pooled OLS estimations

We first estimate the following equation (4) using OLS with the pooled dataset from all of the three survey waves (CFPS 2010, 2012, and 2014).

$$Edu_{ijt} = z_0 + z_1 Oppoineq_{jt} + \gamma X_{ijt} + c_j + y_t + \varepsilon_{ijt} \quad (4).$$

In equation (4), Edu_{ijt} represents the education expenditures for household *i* in county *j* in year *t.*

$Oppoineq_{jt}$ denotes the derived index for inequality of opportunity in county *j* in year *t* calculated

using the method introduced in the previous section. X_{ijt} includes control variables such as family size, household demographic structure, household income, and the householder's health status, years of education and hukou status. c_j and y_t stand for county and year dummy variables.

We employ two alternative measures for household education expenditures, including the natural logarithm value of annual household education expenditures (an absolute measure) and the ratio of annual education expenses to total expenditure (a relative measure). Table 2 reports the results of the pooled OLS, and all estimations include year and county dummy variables. Standard errors are robust to heteroskedasticity and clustered at the county level (hereafter).

Table 2 Inserted Here

The empirical results are consistent with the predictions made by World Bank (2005), Mejía and St-Pierre (2008), and Marrero and Rodriguez (2013). That is, inequality of opportunity has a significantly negative effect on education expenditures no matter which of the two measures we use for the dependent variable. The household demographic variables have expected effects on education expenses. For example, households with fewer children or young people spend less on education. Moreover, if a householder receives more education, the household spends more on education holding other variables constant. Interestingly, higher household income reduces the share of education expenditures as shown in columns (5) and (6).

4.2 Fixed effect estimations

We then estimate equation (4) using the two-way fixed effect panel regression model which can control the effect of time trend and all of the invariant household characteristics with respect to time. Again, the clustered standard errors at the county level are used throughout. We have used the Hausman test to confirm that fixed effect specifications are preferred to random effects specifications for our data.

Table 3 presents the results from the fixed effect model. We find that inequality of opportunity has a more significantly negative effect on both the logarithm value of education expenditures and the ratio of education expenditures to the total expenditure, indicating that higher inequality of opportunity reduces people's incentives for efforts and education investment.

Table 3 Inserted Here

Moreover, we don't find evidence supporting the status seeking hypothesis since the total income inequality of a county (measured by Theil index) has non-positive effects on education expenditures. As Jin et al. (2011) claim, one possibility when the status-seeking theory does not work is that poor people may give up: when inequality is too large, the poor may not be able to invest enough through education to enter the next club, and thus they simply give up. Our empirical findings indicate that the poor in China may give up to spend more on education in the context of high inequality. Another reason might be that poor households cannot afford their children's education even with status-seeking incentives, since they face borrowing constraints in the context of high inequality.

4.3 Heterogeneity analysis

To obtain more understanding of which group is affected most by inequality of opportunity, we conduct a series of heterogeneity analysis by dividing the sample into several sub-groups by different dimensions.

We first divide the whole sample into three subgroups by the household income, and Table 4 informs us that the main negative effect comes from the middle income group. That is, as the inequality of opportunity within a county becomes larger, middle income households may have fewer incentives to invest on education. Poor households are not affected significantly by higher inequality of opportunity in part because they are financially constrained and originally spend little on education. Moreover, as expected, the effect on richer households is insignificant since their incentives to invest in education may not be affected by inequality of opportunity in the county.

The largest impact found for the middle income group is an important finding especially in China's context since China is struggling on avoiding the middle-income trap and hoping to enlarge the share of middle income families. If inequality is more driven by factors beyond people's control, then these middle income families may lose confidence or give up, which might become a serious obstacle for China's future economic development.

Table 4 Inserted Here

We then conduct a similar heterogeneity analysis as above by dividing the whole sample into three subgroups by the householder's education level, and find that the main effect comes from the low and middle education groups.

Table 5 Inserted Here

A final heterogeneity analysis is to divide the whole sample into two subgroups by the householder's *hukou* status (urban and rural). We find in Table 6 that households with rural hukou are negatively affected by inequality of opportunity. In addition, not only the inequality of opportunity has a negative effect on education expenditures for rural hukou households, the total income inequality (measured by Theil index) has a further negative effect on education expenditures. This result can be easily linked to the credit market imperfection literature mentioned above (Galor and Zeira, 1993; Ferreira, 2001; Mej á and St-Pierre, 2008). That is, relatively poor households don't have the means to finance the accumulation of human capital, and, because they are credit constrained, they end up either not investing in human capital or investing very little.

Table 6 Inserted Here

In summary, the disadvantaged households (whose householders with less education, income, and rural hukou) seem to be affected more by inequality of opportunity within the county they live in. Higher inequality of opportunity in the comparison group may reduce their incentives to investment

more on education.

5. Robustness checks

To lend more support to our main results above, we provide several robustness checks next. First, we consider only households with children below 16 years old and drop the households without children, because there is generally very little education investment for adults. The results are reported in Table 7.

Table 7 Inserted Here

By this robustness check, we can find that the main results still hold that inequality of opportunity has a significantly negative effect on education expenditures, and the total income inequality has a non-significant effect. Moreover, by restricting the sample which may be influenced more in terms of education expenditures, we find that the impacts of inequality of opportunity become larger given larger magnitude of the coefficients.

Second, since the education expenditures data is left-censoring at 0, we use the Tobit model as a robustness check with the county level fixed effect which is similar to the model specification in Table 2 (Panel Tobit model only has random effect model). Table 8 reports the estimated coefficients of Tobit model. Again, the main results still hold with the significantly negative signs on inequality of opportunity in each column.

Table 8 Inserted Here

Third, we test whether the extensive margin exists by re-estimating the equation with the dependent variable being a dummy variable for spending a positive amount on education and using the fixed effect panel regression (which means we use the linear probability model). It turns out in Table 9 that the extensive margin exists in the sense some households facing more inequality of opportunity end up with not investing in human capital. Combining the results from Table 7 and Table 9, we can conclude that both the intensive and extensive margins exist in our case.

Table 9 Inserted Here

In summary, all of the robustness checks lend support to our main finding that inequality of opportunity has a significantly negative effect on household education expenditures.

6. Conclusions

This paper explores the relationship between inequality of opportunity and education expenditures by using the micro data from China Family Panel Studies (CFPS). The paper contributes to the literature by offering the first empirical evidence on the impact of inequality of opportunity on household education investment. By using the fixed effect model as the main model specification, our results suggest that even after controlling for total income inequality, inequality of opportunity itself has a

negative effect on household education expenditures. This result is robust to a series of robustness checks.

Furthermore, the disadvantaged households (whose householders with less education, income, and rural hukou) seem to be affected more by inequality of opportunity within the county they live in. Higher inequality of opportunity in the comparison group may reduce their incentives to investment more on education. This is an important finding in a middle-income country like China since China is struggling on avoiding the middle-income trap and hoping to enlarge the ratio of middle income families. If inequality is more driven by factors beyond people's control, then these middle income families may lose confidence or give up, which might become a serious obstacle for China's future economic development.

Given these empirical results, policy actions need to be taken to reduce inequality of opportunity in order to motivate middle-income families to invest more in education and foster economic growth. That is, we have to lower the proportion of income inequality generated by circumstances variables that are beyond people's control. Since the circumstances variables in this paper include gender, household registration status (hukou), paternal and maternal education, policy suggestions may include reducing labor market discrimination based on gender and hukou, balancing education resources to create more equal educational opportunities, and offering children education subsidies in low-income families.

References

Arneson, R., 1989, "Equality and Equal Opportunity of Welfare," Philosophical Studies, 56(1), pp. 77-93.

Alesina, A. and E. La Ferrara (2005), Preferences for redistribution in the land of opportunities, *Journal of Public Economics* 89, 897-931.

Barro, R., 2001. Human capital and growth. *American Economic Review* 91 (2), 12-17.

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

Bourguignon, F., Ferreira, F., Menéndez, M., 2007. Inequality of opportunity in Brazil. *Review of Income and Wealth* 53, 585–618.

Benjamin, D., Brandt, L. and Giles, J. (2011), "Did higher inequality impede growth in rural China", *The Economic Journal*, 121, 1281-1309.

Checchi, D., Peragine, V., 2010. Inequality of opportunity in Italy. *Journal of Economic Inequality* 8, 429–450.

Corneo, Giacomo, Jeanne, Olivier, 1999. Social organization in an endogenous growth model. *International Economic Review* 70, 37 – 51.

Dworkin, Ronald, "What is Equality? Part 2: Equality of Resources," *Philosophy and Public Affairs*, 10(4), 283–345, 1981.

Ferreira, Francisco H. G. (2001): "Education for the Masses? The Interaction between Wealth, Educational and Political Inequalities." *Economics of Transition* 9(2): 533–52.

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

Ferreira, F. H. G., Lakner, C., Lugo, M. A., & Ozler, B. (2014). Inequality of opportunity and economic growth: a cross-country analysis. IZA DP No. 8243.

Galor, Oded and Joseph Zeira (1993): "Income distribution and macroeconomics", *Review of Economic Studies* 60: 35-52.

Hassine, N, 2012. Inequality of Opportunity in Egypt, *World Bank Economic Review*, 26, 265-295.

Jin, Ye, Li, H. and Wu, B. (2011). Income Inequality, Consumption, and Social-Status *Seeking. Journal of Comparative Economics*, 39(2), 191-204.

Kanbur, R., & Wagstaff, A. (2015). Inequality of opportunity: reply to pedro rosa dias and erik schokkaert. *Health Economics, 24*(10), 1253–1255.

Krueger, Alan (2012): "The rise and consequences of inequality in the United States", Remarks delivered to the Center for American Progress, Washington, DC.

Lefranc, A., Pistolesi, N., Trannoy, A., 2008. Inequality of opportunities vs. inequality of outcomes: are Western societies all alike? *Review of Income and Wealth* 54, 513–546.

Lefranc, A., Pistolesi, N., & Trannoy, A. (2009). Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France. *Journal of Public Economics*, 93(11), 1189-1207.

Lucas, R., 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22 (1), 3-42.

Mankiw, G., Romer,D.,Weil,D., 1992. Acontribution to the empirics of economic growth. *Quarterly Journal of Economics* 107 (2), 407-438.

Marrero, G. A., & Rodr ́guez, J. G. (2013). Inequality of opportunity and growth. *Journal of Development Economics*.

Marrero, G. A., & Rodr ́guez, J. G. (2012). Inequality of opportunity in Europe. *Review of Income and Wealth, 58*(4), 597-621.

Mej á, D., & St-Pierre, M. (2008). Unequal opportunities and human capital formation. *Journal of Development Economics*, 86(2), 395-413.

Meng, Xin, Gregory, Robert and Wang, Youjuan (2005). "Poverty, Inequality, and Growth in Urban China, 1986-2000." *Journal of Comparative Economics*, 33, 710-729.

Ramos, X., & Gaer, D. (2015). Approaches to inequality of opportunity: principles, measures and evidence. *Journal of Economic Surveys*, 30 (5), 855-883.

Roemer, J., 1998, Equality of Opportunity, Published by Harvard University Press, Cambridge.

Sen, A., 1985, Commodities and Capabilities, Published by North-Holland, Amsterdam.

Shorrocks, A. F. (1984),"Inequality decomposition by population subgroups, " *Econometrica*, 52,1369-1385.

Song, Yang. (2013). Rising Chinese regional income inequality: the role of fiscal decentralization. *China Economic Review*, 27(27), 294–309.

Song, Yang. (2014). What should economists know about the current Chinese hukou system? *China Economic Review* 29: 200–212.

Song, Yang. (2017). Measures and Underlying mechanisms of Inequality of Opportunities in China: Evidence from the Recent CGSS Data. *Finance & Trade Economics*, 1, 35-51 (in Chinese).

Sun, Wenkai & Xianghong Wang. (2013). Do relative income and income inequality affect consumption? evidence from the villages of rural china. *The Journal of Development Studies*, 49(4), 533-546.

World Bank (2005), World Development Report 2006: Equity and Development. Washington, DC: World Bank.

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

Tables and Figures

Table 1 Summary Statistics of the Variables

Variable	Definition	2010			2012			2014		
		Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
lneduexp	log of household education expenditures	9,201	4.020	3.959	9,138	3.952	4.069	9,227	3.804	4.146
eduratio	education expenditures/total expenditures	9,201	0.105	0.178	9,138	0.097	0.164	9,227	0.073	0.135
theil	theil index for total income inequality	9,274	0.341	0.213	9,274	0.427	0.217	9,274	0.316	0.128
oppoineq	index for inequality of opportunity	9,274	0.173	0.125	9,274	0.240	0.152	9,274	0.224	0.139
family size	family size	9,274	3.886	1.733	9,274	3.812	1.792	9,274	3.660	1.882
childratio	the ratio of children aged below 16	9,274	0.308	0.251	9,274	0.214	0.216	9,274	0.211	0.229
youngratio	the ratio of people aged 16-25	9,274	0.082	0.151	9,274	0.117	0.185	9,274	0.107	0.219
elderratio	the ratio of elder people aged above 60	9,274	0.174	0.305	9,274	0.193	0.321	9,274	0.244	0.356
log hh. income	log of household income	8,820	9.935	1.052	8,336	9.930	1.422	8,501	10.140	1.326
education	education level of the householder	9,269	6.595	4.601	9,272	6.382	4.665	9,274	6.381	4.665
rural hukou	rural hukou status of the householder	9,269	0.271	0.444	9,264	0.291	0.454	9,273	0.295	0.456
health	health status of the householder	9,273	0.820	0.384	9,271	0.794	0.404	9,274	0.809	0.393

Table 2 The Impacts of Inequality of Opportunity on Education Investment (Pooled OLS)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Log Edu Exp.			Edu Exp. Ratio		
theil	-0.1751	-0.3687*	-0.3119	-0.0044	-0.0117*	-0.0150**
	(0.181)	(0.212)	(0.244)	(0.006)	(0.007)	(0.007)
oppoineq		-0.7472**	-0.6920*		-0.0281**	-0.0251**
		(0.363)	(0.424)		(0.012)	(0.012)
family size			0.5508***			0.0089***
			(0.030)			(0.001)
childratio			3.5552***			0.0651***
			(0.206)			(0.008)
youngratio			3.4240***			0.2118***
			(0.295)			(0.014)
elderratio			-1.4631***			-0.0352***
			(0.100)			(0.004)
log hh.income			0.0040			-0.0041***
			(0.026)			(0.001)
education			0.0799***			0.0018***
			(0.008)			(0.000)
rural hukou			0.5033***			-0.0009
			(0.093)			(0.003)
health			0.2428***			0.0040
			(0.067)			(0.003)
Constant	3.7424***	3.8296***	0.2100	0.0845***	0.0878***	0.0603***
	(0.064)	(0.082)	(0.295)	(0.002)	(0.003)	(0.010)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
County Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,566	27,566	25,494	27,566	27,566	25,494
R-squared	0.043	0.043	0.244	0.040	0.041	0.125

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The other variables are self-explanatory, and their definitions have been provided in Table 1. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1

Table 3 The Impacts of Inequality of Opportunity on Education Investment (FE Panel Regression)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables		Log Edu Exp.			Edu Exp. Ratio	
theil	-0.1563	-0.3588*	-0.3511	-0.0036	-0.0107	-0.0134*
	(0.176)	(0.205)	(0.222)	(0.006)	(0.007)	(0.007)
oppoineq		-0.7819**	-0.7787**		-0.0275**	-0.0263**
		(0.352)	(0.392)		(0.012)	(0.012)
family size			0.5675***			0.0134***
			(0.041)			(0.002)
childratio			0.7957***			0.0304***
			(0.186)			(0.008)
youngratio			-0.2408			0.0136
			(0.242)			(0.012)
elderratio			-0.2034			0.0116*
			(0.157)			(0.006)
log hh.income			-0.0149			-0.0048***
			(0.022)			(0.001)
education			-0.0278			-0.0021**
			(0.020)			(0.001)
rural hukou			-0.1360			-0.0032
			(0.182)			(0.008)
health			0.0102			0.0034
			(0.063)			(0.003)
Constant	4.0698***	4.2740***	2.2652***	0.1061***	0.1132***	0.1091***
	(0.070)	(0.124)	(0.333)	(0.002)	(0.004)	(0.013)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Hausman Test	1.17(0.76)	7.33(0.12)	708.64(0.00)	6.58(0.09)	8.28(0.08)	683.20(0.00)
Observations	27,566	27,566	25,494	27,566	27,566	25,494
R-squared	0.002	0.002	0.054	0.017	0.018	0.031
Number of fid10	9,274	9,274	9,261	9,274	9,274	9,261

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 4 Heterogeneous Impacts of IO for Different Income Groups

	(1)	(2)	(3)	(4)	(5)	(6)
		Log Edu Exp.			Edu Exp. Ratio	
Income Groups	low	middle	high	low	middle	high
theil	-0.7272*	-0.3733	-0.1195	-0.0109	-0.0198*	-0.0108
	(0.375)	(0.232)	(0.279)	(0.017)	(0.011)	(0.012)
oppoineq	-0.9710	-1.3616***	-0.2962	-0.0087	-0.0504**	-0.0187
	(0.671)	(0.511)	(0.561)	(0.031)	(0.020)	(0.023)
household control variables	Yes	Yes	Yes	Yes	Yes	Yes
householder's control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,983	8,060	8,088	7,983	8,060	8,088
R-squared	0.075	0.068	0.039	0.026	0.033	0.047
Number of fid10	2,832	2,834	2,833	2,832	2,834	2,833

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 5 Heterogeneous Impacts of IO for Different Education Categories

	(1)	(2)	(3)	(4)	(5)	(6)
		Log Edu Exp.			Edu Exp. Ratio	
Education Categories	illiterate	primary or middle school	high school and above	illiterate	primary or middle school	high school and above
theil	-0.4319	-0.6351**	0.1171	-0.0043	-0.0256***	-0.0042
	(0.363)	(0.261)	(0.224)	(0.012)	(0.008)	(0.013)
oppoineq	-1.2996*	-0.9117**	-0.4278	-0.0156	-0.0389**	-0.0210
	(0.713)	(0.455)	(0.526)	(0.027)	(0.016)	(0.025)
household control variables	Yes	Yes	Yes	Yes	Yes	Yes
householder's control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,261	13,433	4,800	7,261	13,433	4,800
R-squared	0.069	0.054	0.055	0.024	0.032	0.052
Number of fid10	2,660	4,885	1,716	2,660	4,885	1,716

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 6 Heterogeneous Impacts of IO for Different Rural hukou Status

	(1)	(2)	(3)	(4)
	Log Edu Exp.		Edu Exp. Ratio	
Rural hukou Status	rural	urban	rural	urban
theil	-0.6795***	0.0081	-0.0245***	0.0037
	(0.242)	(0.195)	(0.009)	(0.007)
oppoineq	-1.1745*	-0.4109	-0.0464**	-0.0040
	(0.618)	(0.385)	(0.019)	(0.014)
household control variables	Yes	Yes	Yes	Yes
householder's control variables	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
household fixed effect	Yes	Yes	Yes	Yes
Observations	17,809	7,684	17,809	7,684
R-squared	0.065	0.053	0.034	0.039
Number of fid10	6,531	2,729	6,531	2,729

Note: The dependent variables are natural log of household education expenditures (column 1-2) and the ratio of education expenditures in total household expenditures (column 3-4). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 7 Sensitivity Analysis of IO Estimates by Dropping Households with no Children (FE Panel Regression)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Log Edu Exp.			Edu Exp. Ratio		
theil	-0.3040	-0.3163	-0.3240	-0.0076	-0.0118	-0.0117
	(0.218)	(0.232)	(0.229)	(0.009)	(0.010)	(0.010)
oppoineq	-0.7468*	-0.8560*	-0.8487*	-0.0439***	-0.0456***	-0.0453***
	(0.418)	(0.451)	(0.446)	(0.017)	(0.017)	(0.017)
household control variables		Yes	Yes		Yes	Yes
householder's control variables			Yes			Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,524	19,076	19,064	20,524	19,076	19,064
R-squared	0.001	0.029	0.029	0.013	0.019	0.020
Number of fid10	7,966	7,896	7,894	7,966	7,896	7,894

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 8 Sensitivity Analysis of IO Estimates by Using Tobit Model

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Log Edu Exp.			Edu Exp. Ratio		
theil	-0.7036	-0.5904	-0.6409	-0.0252*	-0.0293*	-0.0308**
	(0.434)	(0.499)	(0.483)	(0.013)	(0.016)	(0.015)
oppoineq	-1.3874**	-1.5703*	-1.5287*	-0.0574**	-0.0660**	-0.0648**
	(0.704)	(0.819)	(0.815)	(0.023)	(0.027)	(0.027)
household control variables		Yes	Yes		Yes	Yes
householder's control variables			Yes			Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
County Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,566	25,513	25,494	27,566	25,513	25,494

Note: The dependent variables are natural log of household education expenditures (column 1-3) and the ratio of education expenditures in total household expenditures (column 4-6). "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table 9 Sensitivity Analysis of IO Estimates by Using a Dummy Variable for Education Investment

	(1)	(2)	(3)
theil	-0.0387**	-0.0389**	-0.0403**
	(0.018)	(0.018)	(0.018)
oppoineq	-0.0609*	-0.0655*	-0.0659*
	(0.034)	(0.034)	(0.034)
household control variables		Yes	Yes
householder's control variables			Yes
household fixed effect	Yes	Yes	Yes
Observations	27,822	25,657	25,638
R-squared	0.006	0.062	0.061
Number of fid10	9,274	9,265	9,265

Note: "theil" denotes the total income inequality in a county, and "oppoineq" denote inequality of opporutnity in income in a county. The control variables are the same as in Table 2. Cluster standard errors at county level are presented in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Figure 1 Trends in Total Inequality and Inequality of Opportunities

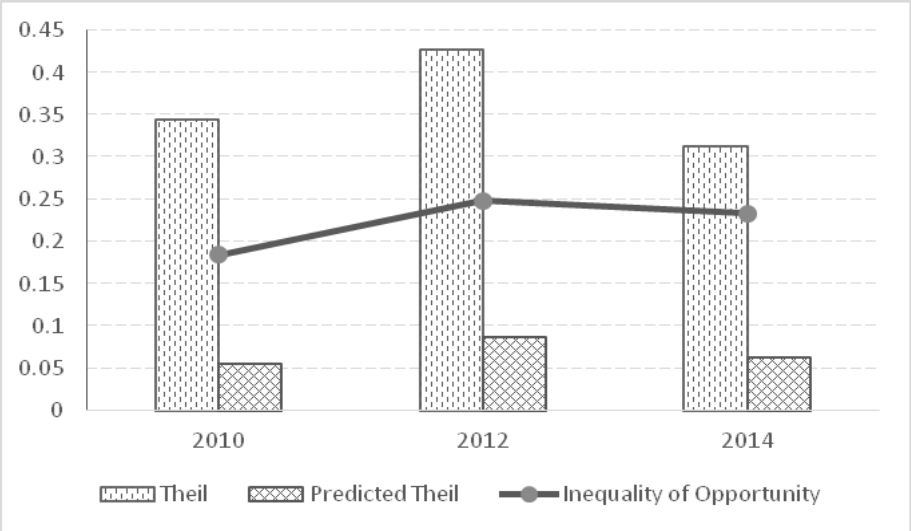
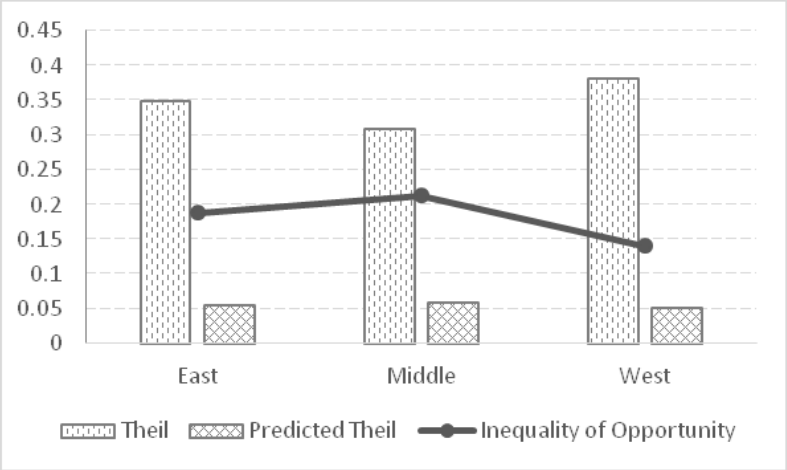
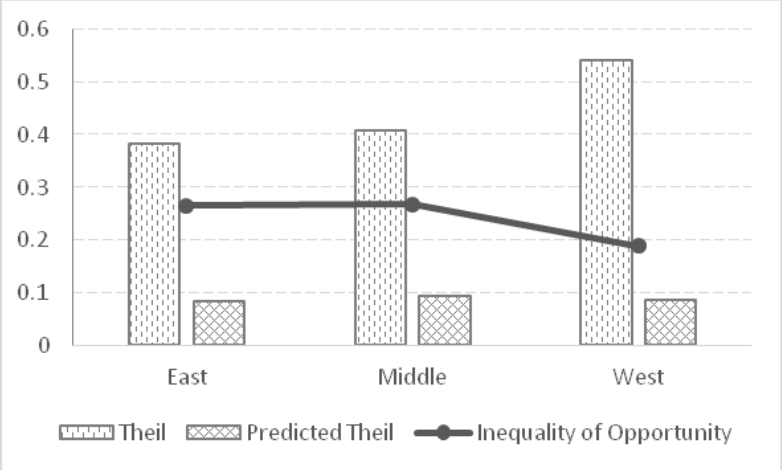


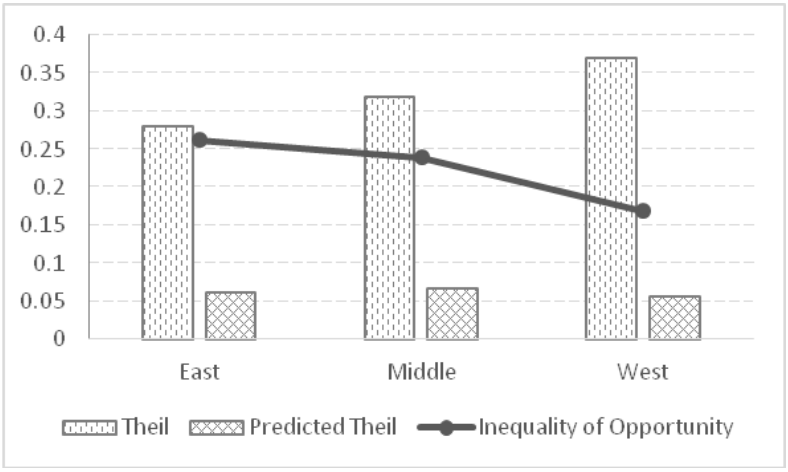
Figure 2 Total Inequality and Inequality of Opportunities in Different Regions



(A) CFPS 2010

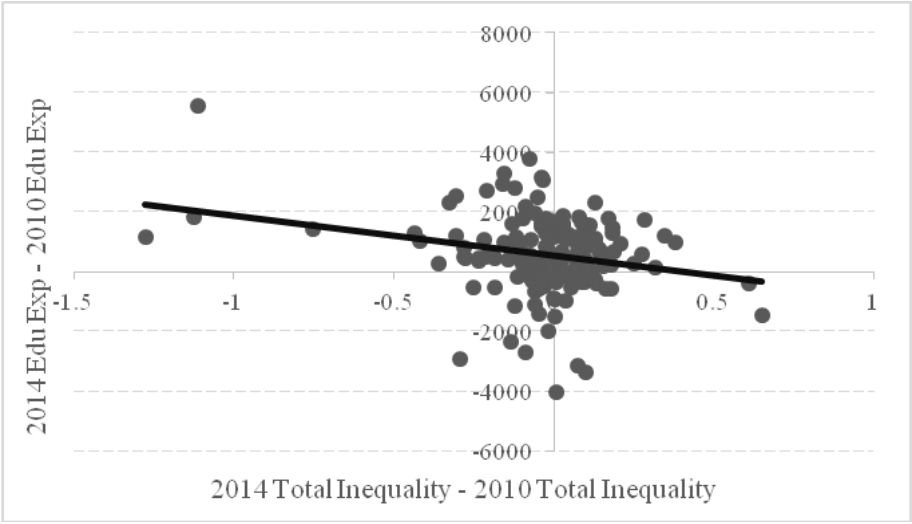


(B) CFPS 2012

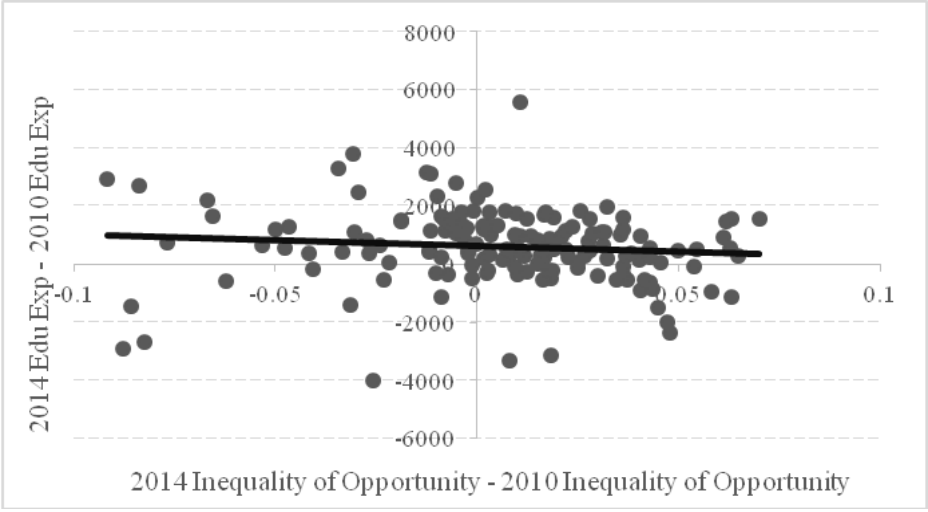


(C) CFPS 2014

Figure 3 Scatter Plot of Total Inequality, Inequality of Opportunities, and Education Expenditure



(A) Total Inequality and Education Expenditure



(B) Inequality of Opportunity and Education Expenditure