How Do Mean Division Shares Affect Development?

Liang Shao

mingliangshao@foxmail.com

School of Economics, Henan University

Jan. 2015

Abstract

Since the seminal paper by Kuznets (1955) there has been an ongoing debate on how the profile of income distribution matters for economic growth. In the present paper we revisit the impact of income inequality on economic development and growth.

This paper studies the growth effects of "mean division shares", i.e. the share of income held by people whose per capita household disposable income is below mean income (mean income share) and the share of population holding this income (mean population share) using a panel data. Thereby, our analysis explores how this income share and population share would impact development and growth. This paper shows that income and population shares affect growth in significantly different ways; thereby providing substantial value added over commonly used summary statistics that aim at compressing the information about the income distribution in a single scalar.

Considering interaction terms of inequality with productive factors driving growth, we also account for different impacts of (the different dimensions of) income inequality on growth depending on the macroeconomic and institutional framework.

In addition to this main contribution of the paper, we estimate and accommodate income inequality by a sole statistic unit and income definition from the data using

different statistic units and income definitions; so that we are able to use more observations than previous studies, by merging several different data sources.

JEL: D31, C36, O11,

I. Introduction

It is still open to debate in academia on how income distribution affects development and growth since Kuznets hypothesis was proposed in 1955. There are two fundamental issues about this problem. The first one is about measurement of (relative) income inequality. Summary index is the most popular measurement of inequality, and economists have found that any summary measurements are not able to strictly rank income distributions, and the correlation between growth and income inequality may differ in inequality measurement. The other question is the mechanics of how income inequality plays its role in production; that is how income inequality should enter a production function. Specifically, functional specification of the macroeconomic effects of income inequality requires to be justified. This paper is motivated to deal with the two issues.

Additionally, data quality is also an important issue for empirical study. There have been collected many data sets of income distribution for many countries by different agents, which had used different statistic units, survey methods and income definitions. It is common that a country's income inequality differs in different data sets due to changes in statistical methods, income definition and sample errors. Fortunately, the team of income distribution projects at WIDER has been working on this issue.

They have professionally examined all previous data sets of income distribution, published all data details, ranked their quality, and made their work publicly available. Their latest release is WIID3b. This paper estimates income inequality from previous data sets using per capita household disposable income based on WIID3b, from which we have retrieved only high quality data and effectively enlarged data size.

In new classic growth theory, concave production function ensures perfect competitive economies to converge toward steady state and thus income inequality eventually would not matter for growth and development. Since we are living in times far away from perfect competition and the so-called "steady state", economists have been presenting theoretical and empirical findings on the correlation between inequality and growth, unfortunately which have not been in line with each other.

Using within-fixed effects regressions, Benhabib and Spiegel (1998), Forbs (2000) and Li and Zou (1998) find a negative relationship between changes in inequality and changes in economic growth rate. Herzer and Vollmer (2012) use heterogeneous panel cointegration techniques and finds that inequality has a negative long run effect on income.

Barro (2000) uses three stage least squares regressions on a panel of countries, and finds no overall relationship between inequality and growth, but he claims there is a negative relationship in the subpanel of poor countries and a positive relationship in the subpanel of rich countries.

Banerjee and Duflo (2003) introduces quadratic form of lagged changes of the Gini index and shows with random effects models that the growth rate is an inverted

U-shape curve of net changes in inequality; changes in the Gini index in any direction are associated with reduced growth in the next period. This is not supported by nonlinearity specification test with dynamic system GMM estimator in our research; and we find an inverted U-shape relationship between output level and the Gini index, and meanwhile a U-shape relationship with a very fat left tail, which is also very similar to a linear negative correlation, between output level and changes of the Gini index.

These researches ignore interactive terms between income inequality and productive factors. The cross items are not described by either country fixed effects or random effects, and they are correlated with both inequality and productive factors; thus missing cross-effect variables make previous research biased and inconsistent.

We also find strong serial correlation in the panel data and thus both fixed effects and random effects models are inappropriate. We will employ robust system GMM estimators with dynamic panel model to deal with endogeneity and heteroskedascity, which gives us consistent estimates. The number of instruments will be properly chosen to deal with the biasedness issue.

Galor and Moav (2004) considers the switch of primary drives of growth from physical capital to human capital, income inequality affects the accumulation of physical capital and human capital and thus affects growth in the two different stages of development. Their hypothesis is roughly supported by the empirical work of Chambers and Clause (2010). This paper shows significant and negative effects for the products of quadratic physical capital stock with income inequality, significant and negative coefficient for the product of quadratic human capital and income inequality;

thus Galor and Moav's hypothesis is strongly supported by this paper's empirical results.

Voitchovsky (2005) finds that the profile of income distribution matters for economic growth. This idea is also supported in this paper by using the mean population share and mean income share, where per capita household income is less than the mean income of an economy, to describe income inequality.

Hernandez. et. al. (2014) finds that it is possible to have rising or falling inequality along with convergent or divergent mobility (changes of income), both in times of economic growth and in times of economic decline. This finding explicitly states that the correlation between inequality and growth is nonlinear. This paper presents strong empirical evidence for Hernandez et. al. (2014).

We present three primary contributions in this paper. Firstly, we use a different measurement of income inequality to explore the macroeconomic effects of inequality. We describe income inequality by the mean population share and mean income share held by the people whose per capita household disposable income is not more than national mean income; which enables us to see how population share and income share play different effects on development and grwoth. We also test explanatory power between the mean division shares and Gini index; it turns out that the mean division shares perform more poweful than the Gini index.

Secondly, we allow inequality to interact with productive factors, which controls for the effects of different structures and institutions in an economy. We run Wald test for the cross items and it shows very strong joint significance in any combinations of the cross items.

Lastly, we find that inequality level and its changes individually and jointly matter for growth and development. The level of inequality and its changes significantly affect growth and development in the next period, but we do not find evidence for an inverted U-shape correlation between growth and the changes in inequality.

II. The Model

2.1 Function specification

We use a simple Cobb-Douglas production function to discuss interaction between inequality and production input factors. Assume there are only two productive factors in the economy, which are per capita labor l and per capita capital stock k. y denotes per capita output, and the production function is $y = l^{\alpha}k^{1-\alpha}$.

Labor share α can be considered as a measurement of income inequality because it denotes the income share of the employed people who rent out labor. If it is assumed that capital holders did not contribute to the labor input l and labor contributors did not hold any capital stock, then l may denote the population share whose people hold the income share α, and the pair values $(l,\ \alpha)$ just describe income inequality of the income distribution in this model economy; of course, wealth inequality is not expressed here. Let's write the production function in log form

$$lny = \alpha lnl - \alpha lnk + lnk \tag{2.1.1}$$

Equation (2.1.1) shows that the log equation of output for a Cobb-douglas technology economy consists in the interactive items between inequality measurements, $\alpha,$ and log forms, lnl and $lnk,$ of input factors.

Assuming that production factors are not strictly exogenous, for instance, employment and investment are affected by wealth distribution. We can have an expression of the marginal macroeconomic effects of income inequality, measured by $\alpha,$ as follows from the equation (2.1.1)

$$\Delta(lny) = \Delta\alpha lnl + \alpha g_l - \Delta\alpha lnk - \alpha g_k \tag{2.1.2}$$

g_l and g_k denote the growth rate of labor and capital stock, respectively. Equation (2.1.2) shows that changes of inequality and input factors play roles for the marginal output effects of inequality. Specifically, there are two types of product between inequality and input factors for the expression of marginal effects of inequality; one is between changes of inequality and log form of input factors, another one is between level of inequality and growth rate, g_l and g_k, of input factors.

For empirical study, we may want to accommodate data from both developed and developing economies to one regression function, then, we can allow all level variables in (2.1.1) and one-period lag of GDP to take quadratic form, and then test the nonlinearity specification to refine the model. This implies that the development level of an economy is denoted by levels of input factors and one-period lag of per capita GDP.

To be concise, we submit the letter of log for all variables. Then let y_t denote the log form of per capita output at period t, x_{jt} the log form of input factor j at period t,

α, β_j, and γ_j parameter to be estimated, m_t is a measurement of income inequality at period t, $f(..)$ is a quadratic function of m_t and m_{t-1}, and g_{jt} is growth rate of factor j at period t, ε_t is the error term.

We apply the following dynamic level output equation (2.1.3), which is generated from equation (2.1.1) at period *t* and *t*-1, to a panel data in this paper.

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_1 {y_{t-1}}^2 + f(m_t, m_{t-1}) + \sum_{p=0,1} \sum_j \left[\beta_{t-p,j} \left(x_{jt-p}\right)^2 m_{t-p} + \gamma_j x_{jt-p}\right] + \varepsilon_t \qquad (2.1.3)$$

To deal with the endogeneity issue of (2.1.3), we use all one-period and after lags of changes of all endogenous and predetermined explanatory variables to instrument the first difference equation, and one-period lag of changes of all explanatory variables to instrument the level equation, and use all exogenous explanatory variables to be standard instruments. This is the one-step system GMM estimator created by Arellano and Bond (1991). We run the estimation on STATA and choose robust option. When the number of instruments is not well chosen, we may have biased but consistent estimates. We will choose the optimal number of instruments by minimizing mean squared error according to Okui (2009).

2.2 Test for nonlinearity specification

Banerjee and Duflo (2003) adds quadratic forms of level inequality and one-period lag of changes in inequality to explain growth, but does not consider interactions between inequality and factor inputs; and they neither test for nonlinearity specification. These jobs are tried to be done in this paper.

We test the nonlinearity specification of equation (2.2.3) in the following way. Firstly we get the fitted value, $\hat{y}$, from (2.2.3), and generate $\hat{y}^2$ and $\hat{y}^3$, secondly put $\hat{y}^2$ and $\hat{y}^3$ in (2.2.3) as explanatory variables to get (2.2.1),

$$y_t = \alpha_0 y_{t-1} + f(m_t, m_{t-1}) + \sum_{p=0,1} \sum_j \left[\beta_{t-p,j} \left(x_{jt-p}\right)^2 m_{t-p} + \gamma_j x_{jt-p}\right] + \hat{y}^2 + \hat{y}^3 + \alpha_1 \varepsilon_t \tag{2.2.1}$$

Thirdly, run (2.2.4) in the same way as (2.2.3), and do Wald test for the joint significance of $\hat{y}^2$ and $\hat{y}^3$. If they are jointly significant, then we reject the nonlinearity specification of (2.2.3); otherwise, we do not reject the specification.

2.3 Function choice

There can be multiple combinations of m_t and m_{t-1} for $f(m_t, m_{t-1})$. For instance, we may also test for Banerjee and Duflo (2003)'s specification, lagged changes in inequality enter the inequality function $f(m_t, m_{t-1}, \Delta m_{t-1}, \Delta m_{t-1})$. So that there can be multiple versions for (2.2.3) due to changes in the inequality function $f(. . .)$; some of which may not be rejected by the above nonlinearity specification test, then we need further rank these functions according to their explanatory power.

We take the following steps to choose a more powerful one between two optional functions that are not rejected by the nonlinearity specification test. Let $\hat{y}_1$ and $\hat{y}_2$ be the fitted value for two optional functions, we put the fitted value of one optional function into the other optional function and run the new regressions with the same method used for the optional functions; if only one of the two new variables of fitted value is significant, then we choose the function that its fitted value is significant in the regression of other optional function; if the two variables of fitted values are both significant (insignificant), then we cannot determine which function is more powerful than the other one.

All regressions will be reported if none is more powerful than the other one.

III. Data

We will use two measurements for income inequality. One is the Gini index and the other one is the mean division point to be introduced in the following subsection 3.1.

3.1 Mean division point of income distribution[1]

Definition: Mean division point (MDP) of a smooth Lorenz curve locates at the point where the slope is unit. The corresponding coordinates are called mean population share and mean income share, respectively.

Mean population share and mean income share are called mean division shares, MDS for short hereafter.

Let $f(w)$ be the probability density function of income distribution, with w denoting income level. Accordingly, $F(w)$ is the cumulative probability function of population share with individual income no more than w, μ is the mean of per capita income in the economy, and (x, y) is a point on the Lorenz curve. Then, we have the following Lorenz curve,

$$y(x) = (1/\mu)\int_0^{F^{-1}(x)} w dF(w) = (1/\mu)\int_0^x F^{-1}(t)dt \qquad (3.1.1)$$

Then, the mean division point (x^*, y^*) on the Lorenz curve is defined by the following equations:

[1] Shao (2010, 2014) has discussed this concept in details.

$$\begin{cases} x^* = F(\mu) \\ y^* = (1/\mu)\int_0^{F(\mu)} F^{-1}(t)dt\,, \end{cases} \tag{3.1.2}$$

Mean division point consists of people whose per capita household income is not more than the national mean income. This point is unique for strictly increasing (smooth) Lorenz curves. For economies with different mean income, they may have the same mean division point. Thus we must notify the corresponding development level when we compare mean division points. It is more straightforward to compare mean division shares for countries with similar per capita GDP.

The Pietra ratio is the difference between mean population share and mean income share. Shao (2010, 2014) shows that the Gini index is about 1.3 times of the Pietra ratio in the panel data we use in this study.

3.2 Data sources

Our macroeconomic data (GDP, population, capital stock, human capital, employment, investment, import and export, government consumption share, and TFP) are retrieved from the PWT8.0 (Penn World Table). The PWT8.0 provides national income accounts with purchasing power parity converted to international prices for 167 countries/territories for some or all of the years from 1950 to 2011.

The data of income distribution are borrowed from “WIID3b” (World Income Inequality Database) of WIDER at the United Nations University. “WIID3b” is a panel data built from different earlier works of income distribution of countries all over the world; these data were collected by different agents, and thus vary in the definition of income, coverage of sample area, household, age, and population. The team claims that the data-set can be comparable without correction and adjustment in definition, statistical unit and survey method; but we find some data using different statistic units and income definitions, thus we collect and adjust data with the following criteria to accommodate different data sources to our regression model.

Income is defined as disposable income [2] measured by per capita household. Sample coverage is nationwide and over all ages of the population. However, if the data available were not sampled by these standards, we will either transfer them by the method discussed in the subsection 3.3, or take an approximation available for a few observations; for instance, disposable income is approximated by the squared root of economic family equivalence for Canada.

"WIID3b" ranks quality of an observation according to its income definition, survey quality, and sample methods. Quality 1 is assigned to those in which "the underlying concepts are known", and "the quality of the income concept and the survey can be judged as sufficient" according to some strict criteria. Quality 2 is assigned to observations that "the quality of either the income concept or the survey is problematic or unknown or we have not been able to verify the estimates (the sources were not available to us)"[3], and all other criteria are satisfied. We choose only the observations rated as quality 1 or quality 2 in this study.

In the data-set there are some observations whose surveyed income was gross (monetary) income or disposable monetary income, or whose income is measured by

---

[2] Its definition refers to the "World Income Inequality Database User Guide", on the table 1 of page 6, and the definition on page 10.

[3] World Income Inequality Database User Guide and Data Sources, page 13. http://website1.wider.unu.edu/wiid/wiid-documentation1.php

household or household equivalence (OECD method, HBAI, square root, see the WIID2 user guide), which can be used to impute the Gini index that agrees with our standards. This enables us to enlarge our dataset as much as possible.

An observation of income distribution will be kept in our dataset even if it does not satisfy our standards, but the Gini index satisfying our standards is available for the year of the observation. In this case the Gini index can be used to estimate the MDS. In the next subsection 3.3, we discuss how to estimate the mean division shares to satisfy our statistical standards.

One country's data may come from different sources to enlarge the data-set size, while one observation for the same year from one source may significantly differ from those of other sources even if they used the same statistical method. When this occurs, we choose the data with more observations and/or the latest version. However, we have to accept that there exist statistical errors in the pool of different data sources, which cannot be completely overcome.

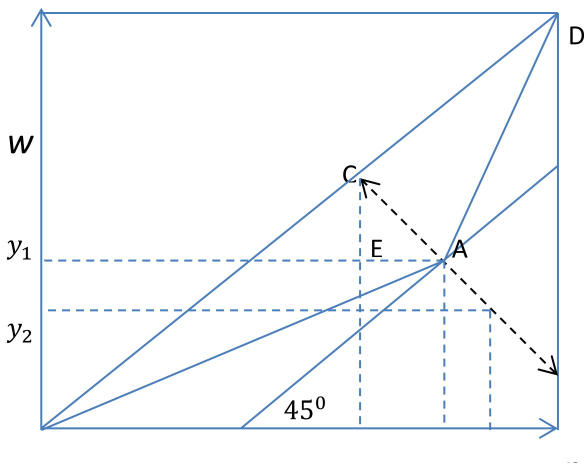
From the two data resources, PWT8.0 and "WIID3b", our analysis is conducted using an unbalanced panel of 31 countries, from 1956 to 20011 using non-overlapping 5-year averages; which are all capable of being cleaned using the above standards and all have the variables required in our models. All macroeconomic variables are measured by per capita household, using log value, and demeaned with overall sample mean.

3.3 Estimation of MDS with Given Gini Index

There are many cases in which a dataset was not made with the methodology we are using in this paper. Then we should find a way to estimate MDS so that it will satisfy our needs. The Lorenz curve changes with measuring units in income and population, and with the definition of income. In these cases, the changed Lorenz curve is unknown, but if the Gini indices are known before and after the change, then the MDS of the changed Lorenz curve can be approximated by the following method.

To simplify the question, we consider only such changes that make the Lorenz curve shift along the orthogonal direction of the tangent line at the MDP. This simplification assumes that the change in measuring unit or income definition proportionally affects each point on the Lorenz curve. Figure 3.2.1 below shows the shifting of a triangle Lorenz curve. Let g_1 be the Gini index of the Lorenz curve $\widehat{OAD}$, and $A(x_1, y_1)$ be its MDP; g_2 is the Gini index and $B(x_2, y_2)$ is the MDP of the Lorenz curve $\widehat{OBD}$ after a change in measuring unit or income definition.

Figure 3.3.1 Shifting of Lorenz Curve



The assumption of a shifting change is not enough to describe point B by point A, but the problem can be resolved if we further approximate the ratio, g, of Gini indices by the squared ratio of heights on the common bottom line $\overline{OD}$ of the two triangles, ΔOAD and ΔOBD. That is, if the two triangles ΔOAD and ΔBDD are assumed to be similar when the shift is very small, then,

$$g_2/g_1 \approx (BC/AC)^2 = g \tag{3.3.1}$$

Employing the assumption (3.3.1), we have the following results

$$\sqrt{g_1/g_2} \approx AC/BC = \left(x_1 - \frac{x_1+y_1}{2}\right)\Big/\left(x_2 - \frac{x_1+y_1}{2}\right)$$
$$= \left(\frac{x_1+y_1}{2} - y_1\right)\Big/\left(\frac{x_1+y_1}{2} - y_2\right)$$
$$\therefore \begin{cases} x_2 \approx .5x_1\left(1+\sqrt{g}\right) + .5y_1\left(1-\sqrt{g}\right) \\ y_2 \approx .5x_1\left(1-\sqrt{g}\right) + .5y_1\left(1+\sqrt{g}\right) \end{cases} \tag{3.3.2}$$

Employing the Gini ratio g and MDP before the shift of Lorenz curve, we can estimate the new MDP after the shift with equations (3.3.2). Fortunately, the Gini index is widely available with different definitions and statistical methods in the data set “WIID3b”, and so the Gini ratio g can be easily computed, which enables us to estimate point B of the shift.

We can directly get the Gini ratio when the Gini indices are available before and after a change of the statistical method; otherwise we choose the average of the Gini ratio of the previous two or three periods when they satisfy our standards.

The transference from household equivalence by the OECD method to household per capita may suffer some errors when we use the equations (3.3.2), but we can only take (3.3.2) as an approximation since we do not have the data of household size for the two population groups.

Finally, we have enlarged our data set using all current income resources, and have kept a uniform statistic unit within each country, but we are still unable to apply the uniform statistic unit to all countries due to lack of information[4]. Data summary is reported below.

Table 3.3.1 Data Summary

Variable	Obs	Mean	Std. Dev.	Min	Max
gdpa	201	4.0891	0.3326	2.8400	4.7461
xa	198	0.6367	0.0664	0.5174	0.8047
ya	198	0.3801	0.0350	0.2813	0.4750
ginia	198	0.3551	0.1092	0.1627	0.6311
hca	194	2.0857	0.3740	0.5967	2.7481
labsha	195	0.5804	0.0999	0.2406	0.7644
empa	197	3.7874	3.0079	-3.9821	14.2799
ma	195	17.0402	3.0642	0.0039	22.8121
ka	197	22.5390	2.1654	13.3418	25.8324
ia	197	16.9227	2.5435	-0.0194	20.8662
xpa	197	16.6423	3.4142	-0.5173	22.2106
gova	197	16.5146	2.1174	0.0558	19.2393

Note: data are 5-year averaged. gdpa is log per capita GDP, xa is the mean population share, ya is the mean income share, ginia is the Gini index, hca is log human capital index, labsha is labor share, empa is employment rate, ma is log ratio of import in GDP, xpa is log ratio of export in GDP, gova is log ratio of government spending in GDP.

IV. Empirical Results

We use two measurements for income inequality, the Gini index and mean division shares in our regressions. We report the results separately in subsection 4.1 and 4.2, and in subsection 4.3 we compare the results of the two measurements and summarize.

4.1 Results using Gini index

[4] There are 20 countries with a total number of 223 observations that do not use the statistic unit of either disposable income or household per capita, which account for about 39% of the observations in the dataset. For instance, income was defined with disposable monetary income for Republic of Korea, Belgium, Switzerland, and Australia, and was defined with gross monetary income for New Zealand and Argentina, and gross income for Honduras, Ukrain, and Uruguay, respectively. The statistic unit of income was square rooted household equivalence for Republic of Korea and Norway, square rooted economic family equivalence for Canada, OECD method household equivalence for Australia, Austria, France, Greece, Ireland, Netherlands, and Portugal, household equivalence (HBAI) for United Kingdom, and tax unit per capita for Switzerland, respectively.

Table 4.1 below reports two regression results for the Gini index. Both have valid nonlinearity specification. Model gdpI has quadratic form of the change in the Gini index and uses robust GMM estimates; model gdpII has linear form of the change in the Gini index and uses the normal GMM estimates because its variance matrix is non-symmetric or highly singular.

We do Wald test for joint significance of the explanatory variables involving the Gini index. For the two functions, we find that all the 12 items involving the Gini index, the 8 items of cross effects between the Gini index and investment, capital stock, human capital, and employment, the 4 items of current or lagged cross items, the 4 items of quadratic Gini index and quadratic change of the Gini index, and the quadratic two items of the Gini index all are jointly significant at 1% significance level; For the model gdpI, the quadratic form of changes in the Gini index are jointly significant at 5% level. For the model I, the ginisqr and deltagini are jointly significant at 1% level. The estimates for other variables are available in the appendix.

Therefore, we find that the Gini index and its change, jointly and individually, present significant effects for development and growth in the same period, and the Gini index and input factors (investment, capital stock, human capital, and employment rate) show significant interactive effects in the same period and next period.

Table 4.1 Robust system GMM estimation with the Gini index

Variable	gdpI	gdpII
gdpa		
L1.	0.4061***	0.3895***
gdpsqr		
L1.	0.2111***	0.2250***
ginia	0.2503***	0.2570***
ginisqr	-0.8243**	-0.7001
deltagini	-0.2837*	-0.2803**
deltaginisqr	1.3488	
isqrgini		
--.	-0.0809***	-0.0831***
L1.	0.0020***	0.0020*
ksqrgini		
--.	0.1191***	0.1393***
L1.	-0.0477	-0.0633*
hcsqrgini		
--.	-0.6284**	-0.7924*
L1.	0.5581**	0.7121**
empsqrgini		
--.	0.0004	-0.0015
L1.	-0.0082	-0.0069

Note: Estimates of input factors are not reported here, detailed regression results can be found in the appendix. deltagini is the change of the Gini index. gdpsqr=(gdpa-4.0891)^2, deltaginisqr=deltagini^2, ginisqr=(ginia-.3551)^2, isqrgini=(ginia-.3551)*(ia-16.9227)^2, empsqrgini=(ginia-.3551)*(empa-3.7874)^2, hcsqrgini=(ginia-.3551)*(hca-2.0857)^2, ksqrgini=(ginia-.3551)*(ka-22.539)^2 .

We tried all possible forms for the inequality function, including the quadratic change of one-period lagged Gini index, which is used by Banerjee and Duflo (2003), the test for nonlinearity specification for all other options suggests rejection of all the other options's validity. Especially, we do not find a valid nonlinearity specification of one-period lag changes of the Gini index according to robust system GMM dynamic estimator. But we do find valid nonlinearity specification of one-period lag changes of mean division shares, which are reported in the next subsection 4.2.

4.2 Results using mean division shares

Table 4.2 below reports the robust system GMM estimates for the model gdpmI with mean division shares, where we only present the estimates for mean division shares.

The full regression results can be found in appendix. We also find another valid nonlinearity specification gdpmII reported in the appendix. gdpmI performs better in explanatory power than gdpmII. If we drop any forms of changes in MDS, then no any nonlinearity specification of level MDS has been found to be valid.

Table 4.2 Robust system GMM estimates for gdpmI

```
Dynamic panel-data estimation                Number of obs        =        87
Group variable: id                           Number of groups     =        31
Time variable: t
                                             Obs per group:    min =         1
                                                               avg =  2.806452
                                                               max =         7

Number of instruments =    121               Wald chi2(29)        =  1.56e+08
                                             Prob > chi2          =    0.0000
One-step results
                                     (Std. Err. adjusted for clustering on id)
```

gdpa	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
gdpa L1.	.2487349	.0556177	4.47	0.000	.1397262	.3577437
gdpsqr L1.	.2059239	.0572785	3.60	0.000	.0936602	.3181877
xa L1.	.6346469	.1206807	5.26	0.000	.398117	.8711768
xsqr L1.	-3.28153	.7747007	-4.24	0.000	-4.799916	-1.763145
ya L1.	-.3587531	.1909546	-1.88	0.060	-.7330171	.015511
deltay L1.	.0128096	.0945708	0.14	0.892	-.1725457	.1981649
deltax	.7541242	.2197328	3.43	0.001	.3234558	1.184793

We do not find nonlinear relationship between log per capita GDP and changes in mean division shares, but mean population share shows an optimal position to induce highest growth in the next period, which is similar to the Gini index.

The Wald tests show that one- period lagged mean income share and its change are jointly significant at 1% level even though each is individually insignificant; change of mean population share and one-period lagged mean income share are also jointly significant at 1% level; interactions between mean division shares and human capital, employment are individually insignificant, but they are jointly significant at 1% level.

We have the following three important findings for mean division shares,

1) Growth rate shows an inverted U-shape with respect to last period's mean population share; thus there is an optimal local mean population share that results highest growth in the next period.
2) Growth is positively and significantly correlated with an increase of mean population share in the same period.
3) Mean income share is negatively correlated with growth in the next period, but an increase of mean income share may enhance growth in the next period;

This model reported in Table 4.2 is called gdpmI. We also find another valid nonlinearity specification, called gdpmII in the appendix, with mean division shares, but this one performs more powerful and the estimates of the two specifications are consistent.

Lastly, we compare the model with the Gini index (gdpI) and the model with MDS (gdpmI). We add the variable of fitted value of one model to the other model and run the new model with the same estimation method, it turns out the over-identification is valid and first order serial correlation is significant as well, and the estimate of fitted gdpI is insignificant, and the estimate of fitted gdpmI is significant, thus we conclude that the model gdpmI performs better than gdpI in explanatory power.

4.3 Summary

According to the robust system GMM estimator on one-period dynamic panel model, both the Gini index and mean division shares give us the following empirics,

1) The level of income inequality shows nonlinear relationship (inverted U-shape) with growth, and changes in income inequality seem to have linear relationship with growth.
2) Income inequality has significant interactive effects, individually and jointly, with productive factors on growth.
3) We do not find any nonlinear relationship between growth and changes in inequality.

We also find some interesting differences between the Gini index and mean division shares about their growth effects.

1) The change of mean population share is significantly and positively correlated with growth in the next period, but the change of the Gini index seems to be negatively correlated with growth in the same period.
2) The level and change of mean income share are individually and jointly insignificant for growth in next period, but the current change in mean population share and last period's change in mean income share are jointly significant for growth.
3) A function with mean division shares can perform better than one with the Gini index.

V. Concluding Remarks

We apply robust system GMM estimator to a one-period dynamic panel model that is heterogenous and serially correlated, and use the same method to test nonlinearity specification for function forms of income inequality; meanwhile we measure income inequality by mean division shares and the Gini index, and allow interactive terms between income inequality and productive factors in the dynamic model. Then, we find that income inequality shows significant nonlinear effects on development and growth, which reply on the levels of income inequality and production factors as well.

We also find that inequality measurement matters for its macroeconomic effects. This paper shows that mean division point performs better than the Gini index according to robust system GMM estimator.

The optimal number of instruments, specification test with instruments estimator, and choice between different measurements of income inequality are correlated issues for the dynamic model. This paper requires further work on these issues to reevaluate our results.

References

1. Arellano, M., and Bond, S.: Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58: 277-297 (1991)
2. Barro, R.J.: Inequality and growth in a panel of countries. J. Econ. Growth **5**, 5–32 (2000)

3. Chambers, D., Krause, A.: Is the relationship between inequality and growth affected by physical and human capital accumulation? J. Econ. Inequal. doi:10.1007/s10888-009-9111-x. (2010)
4. Banerjee, A.V., Duflo, E.: Inequality and growth: what can the data Say? J. Econ. Growth **8**, 267–299 (2003)
5. Forbes, K.J.: A reassessment of the relationship between inequality and growth. Am. Econ. Rev. **90**, 867–889 (2000)
6. Galor, O., Moav, O.: From physical to human capital accumulation: inequality and the process of development. Rev. Econ. Stud. **71**, 1001–1026 (2004)
7. Hernandez, R., D., Fields, G.,S., and Jakubson, G., H.: Changing income inequality and panel income changes in times of economic growth and economic decline. May, 2014
8. Herzer, D. and Vollmer, S., Inequality and Growth: evidence from panel cointegration. Journal Economic Inequality (2012), 10:489–503.
9. Okui, Ryo,: The optimal choice of instruments in dynamic panel models. Journal of Econometrics. 151 (2009) 1-16
10. Shao, M.L.: The Gini index, Pietra ratio and Mean division Shares. Henan University working paper (2014)
11. Shao, M.L.: Two essays on the correlation between growth and income inequality. University of Miami dissertation (2010)
12. Voitchovsky, S.: Does the profile of income inequality matter for economic growth? J. Econ. Growth 10, 273–296 (2005)

Appendix

Table A1 Estimation for the model gdpI with the Gini index

```
Dynamic panel-data estimation                 Number of obs         =        132
Group variable: id                            Number of groups      =         47
Time variable: t
                                              Obs per group:    min =          1
                                                                avg =   2.808511
                                                                max =          8

Number of instruments =    131                Wald chi2(24)         = 154381.86
                                              Prob > chi2           =    0.0000
One-step results
                                        (Std. Err. adjusted for clustering on id)
```

gdpa	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
gdpa						
L1.	.4060743	.0556001	7.30	0.000	.2971	.5150485
gdpsqr						
L1.	.2110566	.0436407	4.84	0.000	.1255225	.2965907
ginia	.2502594	.0728106	3.44	0.001	.1075532	.3929656
ginisqr	-.8243031	.2937137	-2.81	0.005	-1.399971	-.2486349
deltagini	-.2836528	.1220062	-2.32	0.020	-.5227806	-.0445251
deltaginisqr	1.348812	1.041824	1.29	0.195	-.6931249	3.390749
isqrgini						
--.	-.0808894	.0073393	-11.02	0.000	-.0952742	-.0665045
L1.	.0020354	.0004062	5.01	0.000	.0012393	.0028315
ksqrgini						
--.	.1190918	.0264867	4.50	0.000	.0671788	.1710048
L1.	-.0476742	.0256951	-1.86	0.064	-.0980356	.0026872
hcsqrgini						
--.	-.6283724	.2341776	-2.68	0.007	-1.087352	-.1693928
L1.	.5581479	.2096148	2.66	0.008	.1473104	.9689854
empsqrgini						
--.	.0003604	.008257	0.04	0.965	-.015823	.0165438
L1.	-.0081946	.0060533	-1.35	0.176	-.0200588	.0036696
ia	.0310159	.0057589	5.39	0.000	.0197287	.0423031
gova	.0357492	.0085379	4.19	0.000	.0190151	.0524833
empa	.005778	.0010921	5.29	0.000	.0036374	.0079186
xpa	.0148313	.0073278	2.02	0.043	.0004691	.0291935
ka	.0160352	.0072052	2.23	0.026	.0019133	.0301572
ma	.0041624	.0065857	0.63	0.527	-.0087454	.0170702
hca	.0202428	.0170583	1.19	0.235	-.0131909	.0536766
labemp	.0461213	.0151379	3.05	0.002	.0164516	.0757909
labk	-.05201	.0339943	-1.53	0.126	-.1186376	.0146176
labhc	-.0081856	.1590212	-0.05	0.959	-.3198614	.3034903
_cons	.4959472	.0918012	5.40	0.000	.3160201	.6758744

```
Instruments for differenced equation
        GMM-type: L(2/.).gdpa L(1/.).L.gdpsqr L(1/.).ginia L(1/.).ginisqr
                  L(1/.).deltaginisqr L(1/.).deltagini L(1/.).L.hcsqrgini
                  L(1/.).L.isqrgini L(1/.).L.ksqrgini L(1/.).L.empsqrgini
        Standard: D.labhc D.labemp D.labk D.ka D.ia D.gova D.hca D.xpa
                  D.ma D.empa
Instruments for level equation
        GMM-type: L2D.gdpa L2D.gdpsqr LD.ginia LD.ginisqr LD.deltaginisqr
                  LD.deltagini LD.hcsqrgini L2D.hcsqrgini LD.isqrgini
                  L2D.isqrgini LD.ksqrgini L2D.ksqrgini LD.empsqrgini
                  L2D.empsqrgini
        Standard: _cons
```

Table A2 Estimation for the model gdpII with the Gini index

```
Dynamic panel-data estimation                Number of obs         =       132
Group variable: id                           Number of groups      =        47
Time variable: t
                                             Obs per group:    min =         1
                                                               avg =  2.808511
                                                               max =         8

Number of instruments =    127               Wald chi2(23)         =   9749.36
                                             Prob > chi2           =    0.0000
One-step results
```

gdpa	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
gdpa						
L1.	.3895265	.0526557	7.40	0.000	.2863231	.4927298
gdpsqr						
L1.	.2249695	.0419819	5.36	0.000	.1426865	.3072525
ginia	.2570326	.0652683	3.94	0.000	.1291091	.384956
ginisqr	-.7000763	.3761606	-1.86	0.063	-1.437337	.037185
deltagini	-.2803301	.1020712	-2.75	0.006	-.480386	-.0802743
isqrgini						
--.	-.0831321	.0061432	-13.53	0.000	-.0951727	-.0710916
L1.	.0019724	.0009208	2.14	0.032	.0001676	.0037771
ksqrgini						
--.	.1393052	.0320309	4.35	0.000	.0765258	.2020847
L1.	-.0633083	.028592	-2.21	0.027	-.1193477	-.0072689
hcsqrgini						
--.	-.7924173	.3163881	-2.50	0.012	-1.412527	-.172308
L1.	.7121098	.2289809	3.11	0.002	.2633154	1.160904
empsqrgini						
--.	-.0015467	.0089211	-0.17	0.862	-.0190317	.0159382
L1.	-.006898	.007783	-0.89	0.375	-.0221523	.0083564
ia	.0325845	.0052126	6.25	0.000	.0223681	.0428009
gova	.0376454	.0062466	6.03	0.000	.0254023	.0498884
empa	.0055992	.0015677	3.57	0.000	.0025265	.0086718
xpa	.0136878	.0062399	2.19	0.028	.0014579	.0259178
ka	.0176057	.006728	2.62	0.009	.004419	.0307924
ma	.0043169	.0071147	0.61	0.544	-.0096277	.0182614
hca	.0202891	.0183113	1.11	0.268	-.0156003	.0561785
labemp	.0458724	.0178355	2.57	0.010	.0109155	.0808293
labk	-.0554588	.0416366	-1.33	0.183	-.137065	.0261474
labhc	-.0374887	.2033087	-0.18	0.854	-.4359664	.3609891
_cons	.4830943	.0769443	6.28	0.000	.3322862	.6339025

```
Instruments for differenced equation
        GMM-type: L(2/.).gdpa L(1/.).L.gdpsqr L(1/.).ginia L(1/.).ginisqr
                  L(1/.).deltagini L(1/.).L.hcsqrgini L(1/.).L.isqrgini
                  L(1/.).L.ksqrgini L(1/.).L.empsqrgini
        Standard: D.labhc D.labemp D.labk D.ka D.ia D.gova D.hca D.xpa
                  D.ma D.empa
Instruments for level equation
        GMM-type: L2D.gdpa L2D.gdpsqr LD.ginia LD.ginisqr LD.deltagini
                  LD.hcsqrgini L2D.hcsqrgini LD.isqrgini L2D.isqrgini
                  LD.ksqrgini L2D.ksqrgini LD.empsqrgini L2D.empsqrgini
        Standard: _cons
```

Table A3 Estimates using the Gini index

Variable	gdpI	gdpII
gdpa		
L1.	0.4061***	0.3895***
gdpsqr		
L1.	0.2111***	0.2250***
ginia	0.2503***	0.2570***
ginisqr	-0.8243**	-0.7001
deltagini	-0.2837*	-0.2803**
deltaginisqr	1.3488	
isqrgini		
--.	-0.0809***	-0.0831***
L1.	0.0020***	0.0020*
ksqrgini		
--.	0.1191***	0.1393***
L1.	-0.0477	-0.0633*
hcsqrgini		
--.	-0.6284**	-0.7924*
L1.	0.5581**	0.7121**
empsqrgini		
--.	0.0004	-0.0015
L1.	-0.0082	-0.0069
ia	0.0310***	0.0326***
gova	0.0357***	0.0376***
empa	0.0058***	0.0056***
xpa	0.0148*	0.0137*
ka	0.0160*	0.0176**
ma	0.0042	0.0043
hca	0.0202	0.0203
labemp	0.0461**	0.0459*
labk	-0.0520	-0.0555
labhc	-0.0082	-0.0375
_cons	0.4959***	0.4831***
N	132	132
r2		
r2_a		

legend: * p<0.05; ** p<0.01; *** p<0.001

Table A4 Estimation for the model gdpmI with MDS

```
Dynamic panel-data estimation                   Number of obs        =        87
Group variable: id                              Number of groups     =        31
Time variable: t
                                                Obs per group:    min =         1
                                                                  avg =  2.806452
                                                                  max =         7

Number of instruments =    121                  Wald chi2(29)        =  1.56e+08
                                                Prob > chi2          =    0.0000
One-step results
                                      (Std. Err. adjusted for clustering on id)
```

gdpa	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
gdpa						
L1.	.2487349	.0556177	4.47	0.000	.1397262	.3577437
gdpsqr						
L1.	.2059239	.0572785	3.60	0.000	.0936602	.3181877
xa						
L1.	.6346469	.1206807	5.26	0.000	.398117	.8711768
xsqr						
L1.	-3.28153	.7747007	-4.24	0.000	-4.799916	-1.763145
ya						
L1.	-.3587531	.1909546	-1.88	0.060	-.7330171	.015511
deltay						
L1.	.0128096	.0945708	0.14	0.892	-.1725457	.1981649
deltax	.7541242	.2197328	3.43	0.001	.3234558	1.184793
isqrx						
--.	-.1814823	.0144893	-12.53	0.000	-.2098808	-.1530837
L1.	.0038873	.0146617	0.27	0.791	-.0248491	.0326237
isqry						
--.	.2386084	.0516699	4.62	0.000	.1373373	.3398796
L1.	-.0149724	.0505503	-0.30	0.767	-.1140492	.0841044
ksqrx						
--.	.1482767	.0494443	3.00	0.003	.0513676	.2451858
L1.	-.0261274	.0561953	-0.46	0.642	-.1362682	.0840134
ksqry						
--.	-.2431177	.050912	-4.78	0.000	-.3429034	-.1433321
L1.	.1155369	.054006	2.14	0.032	.0096871	.2213868
hcsqry						
--.	-.5035007	1.195559	-0.42	0.674	-2.846754	1.839752
L1.	-1.216172	.9499585	-1.28	0.200	-3.078056	.6457128
hcsqrx						
--.	-.1143459	.6831984	-0.17	0.867	-1.45339	1.224698
L1.	.4882423	.7539054	0.65	0.517	-.989385	1.96587
empsqrx						
--.	-.0044897	.0148208	-0.30	0.762	-.033538	.0245586
L1.	-.0005422	.0158802	-0.03	0.973	-.0316669	.0305825
empsqry						
--.	.0055662	.0157972	0.35	0.725	-.0253958	.0365282
L1.	.0091657	.0098517	0.93	0.352	-.0101433	.0284748
labhc	.0648102	.3365392	0.19	0.847	-.5947946	.7244151
labk	-.0950647	.0454455	-2.09	0.036	-.1841363	-.0059931
labemp	.0553034	.0184868	2.99	0.003	.0190698	.0915369
ka	.0280715	.0069891	4.02	0.000	.0143732	.0417698
ia	.0306795	.0043771	7.01	0.000	.0221006	.0392585
gova	.0504357	.0075903	6.64	0.000	.0355589	.0653125
hca	.0319257	.0321608	0.99	0.321	-.0311082	.0949597
xpa	.0134481	.0082809	1.62	0.104	-.0027821	.0296783
ma	.0083552	.0076372	1.09	0.274	-.0066135	.0233239
empa	.0090032	.0016907	5.33	0.000	.0056895	.0123169
_cons	.3552457	.162973	2.18	0.029	.0358245	.6746668

```
Instruments for differenced equation
        GMM-type: L(2/.).gdpa L(1/.).L.gdpsqr L(1/.).L.xa L(1/.).L.xsqr
                  L(1/.).L.ya L(1/.).L.deltay L(1/.).deltax L(1/.).L.isqrx
                  L(1/.).L.hcsqrx L(1/.).L.ksqrx L(1/.).L.empsqrx
                  L(1/.).L.isqry L(1/.).L.hcsqry L(1/.).L.ksqry
                  L(1/.).L.empsqry
        Standard: D.labhc D.labk D.labemp D.ka D.ia D.gova D.hca D.xpa
                  D.ma D.empa
Instruments for level equation
        GMM-type: L2D.gdpa L2D.gdpsqr L2D.xa L2D.xsqr L2D.ya L2D.deltay
                  LD.deltax LD.isqrx L2D.isqrx LD.hcsqrx L2D.hcsqrx
                  LD.ksqrx L2D.ksqrx LD.empsqrx L2D.empsqrx LD.isqry
                  L2D.isqry LD.hcsqry L2D.hcsqry LD.ksqry L2D.ksqry
                  LD.empsqry L2D.empsqry
        Standard: _cons
```

Table A5 Estimation for the model gdpmII with MDS

```
Dynamic panel-data estimation                  Number of obs      =        87
Group variable: id                             Number of groups   =        31
Time variable: t
                                               Obs per group:    min =         1
                                                                 avg =  2.806452
                                                                 max =         7

Number of instruments =    124                 Wald chi2(30)      =  9.91e+07
                                               Prob > chi2        =    0.0000
One-step results
                                       (Std. Err. adjusted for clustering on id)
```

gdpa	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
gdpa						
L1.	.3310127	.0498938	6.63	0.000	.2332227	.4288027
gdpsqr						
L1.	.2092764	.052287	4.00	0.000	.1067957	.3117571
deltay						
L1.	.0557951	.1082256	0.52	0.606	-.1563231	.2679132
xa	.5809047	.1079153	5.38	0.000	.3693947	.7924147
xsqr	-3.679948	.7514057	-4.90	0.000	-5.152676	-2.20722
ya	-.2876867	.216558	-1.33	0.184	-.7121326	.1367593
deltax						
L1.	-.0887422	.1229807	-0.72	0.471	-.3297799	.1522955
isqrx						
--.	-.1756394	.0151307	-11.61	0.000	-.205295	-.1459838
L1.	.029464	.0109256	2.70	0.007	.0080502	.0508778
isqry						
--.	.2551206	.056687	4.50	0.000	.1440162	.366225
L1.	-.0975564	.0382811	-2.55	0.011	-.172586	-.0225267
ksqry						
--.	-.206176	.0550141	-3.75	0.000	-.3140017	-.0983504
L1.	.1105406	.0563198	1.96	0.050	.0001559	.2209254
ksqrx						
--.	.1316702	.0428353	3.07	0.002	.0477145	.2156258
L1.	-.0250563	.0464777	-0.54	0.590	-.1161509	.0660383
hcsqry						
--.	-.0311214	1.297718	-0.02	0.981	-2.574602	2.512359
L1.	-.7009681	1.132001	-0.62	0.536	-2.91965	1.517714
hcsqrx						
--.	-.3873061	.6783989	-0.57	0.568	-1.716943	.9423313
L1.	.9410841	.82868	1.14	0.256	-.6830989	2.565267
empsqrx						
--.	.0059697	.0136518	0.44	0.662	-.0207873	.0327268
L1.	-.0122655	.0132053	-0.93	0.353	-.0381475	.0136165
empsqry						
--.	-.007283	.0148794	-0.49	0.625	-.036446	.02188
L1.	.0177208	.0095784	1.85	0.064	-.0010526	.0364942
labhc	-.1813435	.3212022	-0.56	0.572	-.8108882	.4482012
labk	-.0644943	.0356614	-1.81	0.071	-.1343894	.0054008
labemp	.0431827	.0163241	2.65	0.008	.0111881	.0751773
ka	.0241822	.0068072	3.55	0.000	.0108404	.037524
ia	.0281342	.0047346	5.94	0.000	.0188546	.0374138
gova	.0440912	.0058096	7.59	0.000	.0327046	.0554779
hca	.0424184	.0313317	1.35	0.176	-.0189907	.1038274
xpa	.011047	.0090297	1.22	0.221	-.006651	.0287449
ma	.0093786	.0081196	1.16	0.248	-.0065356	.0252927
empa	.0083932	.0015487	5.42	0.000	.0053578	.0114286
_cons	.2700822	.2122655	1.27	0.203	-.1459504	.6861149

```
Instruments for differenced equation
        GMM-type: L(2/.).gdpa L(1/.).L.gdpsqr L(1/.).L.deltax
                  L(1/.).L.deltay L(1/.).xa L(1/.).xsqr L(1/.).ya
                  L(1/.).L.isqrx L(1/.).L.hcsqrx L(1/.).L.ksqrx
                  L(1/.).L.empsqrx L(1/.).L.isqry L(1/.).L.hcsqry
                  L(1/.).L.ksqry L(1/.).L.empsqry
        Standard: D.labhc D.labk D.labemp D.ka D.ia D.gova D.hca D.xpa
                  D.ma D.empa
Instruments for level equation
        GMM-type: L2D.gdpa L2D.gdpsqr L2D.deltax L2D.deltay LD.xa LD.xsqr
                  LD.ya LD.isqrx L2D.isqrx LD.hcsqrx L2D.hcsqrx LD.ksqrx
                  L2D.ksqrx LD.empsqrx L2D.empsqrx LD.isqry L2D.isqry
                  LD.hcsqry L2D.hcsqry LD.ksqry L2D.ksqry LD.empsqry
                  L2D.empsqry
        Standard: _cons
```

Table A6 Estimation with MDS

Variable	gdpmI	gdpmII
gdpa		
L1.	0.2487***	0.3310***
gdpsqr		
L1.	0.2059***	0.2093***
xa		
L1.	0.6346***	
--.		0.5809***
xsqr		
L1.	-3.2815***	
--.		-3.6799***
ya		
L1.	-0.3588	
--.		-0.2877
deltay		
L1.	0.0128	0.0558
deltax	0.7541***	
isqrx		
--.	-0.1815***	-0.1756***
L1.	0.0039	0.0295**
isqry		
--.	0.2386***	0.2551***
L1.	-0.0150	-0.0976*
ksqrx		
--.	0.1483**	0.1317**
L1.	-0.0261	-0.0251
ksqry		
--.	-0.2431***	-0.2062***
L1.	0.1155*	0.1105*
hcsqry		
--.	-0.5035	-0.0311
L1.	-1.2162	-0.7010
hcsqrx		
--.	-0.1143	-0.3873
L1.	0.4882	0.9411
empsqrx		
--.	-0.0045	0.0060
L1.	-0.0005	-0.0123
empsqry		
--.	0.0056	-0.0073
L1.	0.0092	0.0177
labhc	0.0648	-0.1813
labk	-0.0951*	-0.0645
labemp	0.0553**	0.0432**
ka	0.0281***	0.0242***
ia	0.0307***	0.0281***
gova	0.0504***	0.0441***
hca	0.0319	0.0424
xpa	0.0134	0.0110
ma	0.0084	0.0094
empa	0.0090***	0.0084***
deltax		
L1.		-0.0887
_cons	0.3552*	0.2701
N	87	87
r2		
r2_a		

legend: * p<0.05; ** p<0.01; *** p<0.001