Economics Bulletin

Volume 41, Issue 2

Proo-poor growth modeling in developing countries: A Gini regression approach

Ndéné Ka
Université Alioune Diop de Bambey

Abstract

The present study analyzes the relationship between poverty, growth, and income inequality in 24 developing countries. It takes a new empirical approach to analyze the effects of growth and inequality on poverty reduction. Mobilizing an unbalanced panel data over the period 2000–2007 and using the Gini fixed effect estimator to address endogeneity issues, selection bias, and error measurement the results suggest that income inequality and inflation negatively impact poverty reduction and there is a positive and robust relationship between growth and the logarithm of agricultural GDP per capital on poverty. Also, the results indicate that Gini methodology addressing the econometric shortcomings of OLS regression analysis may yield more precise results.

Citation: Ndéné Ka, (2021) "Proo-poor growth modeling in developing countries: A Gini regression approach", *Economics Bulletin* , Vol. 41 No. 2 pp. 316-327.
Contact: Ndéné Ka - ndene.ka@uadb.edu.sn.
Submitted: September 26, 2020. **Published:** April 09, 2021.

1. Introduction

For a long time, income inequality has been considered a temporary problem that would disappear naturally with the process of economic growth. Consequently, there was no question of pursuing costly social policies that might lead to fiscal distortions; for this, one could only rely on capital accumulation. Kuznets is certainly one of the first authors to formalize this idea. Already in the mid-1950s, he argued that inequality and growth have a relationship in the form of an inverted "U" shaped function. Indeed, according to him, inequalities initially increase with the growth process. Thereafter, they fall back with economic development. However, far from observing this phenomenon, inequalities continue to persist despite periods of strong economic growth Kai et al. (2009).

Debates since the early 1990s on possible alternatives to the inequality literature have brought a new perspective to the concept of pro-poor growth (or inclusive growth). This new concept, supported by authors such as Anand and Kanbur (1993), Bourguignon (2003), Klasen (2005), discusses the conditions under which growth benefits the poorest. As Klasen (2005) points out, this approach in no way attempts to refute the idea that growth ultimately reduces poverty. Promoting pro-poor growth therefore means giving priority to policies that have a favorable impact on both growth and inequality reduction. Historically, two approaches are considered to address this theme: in the first, so-called relative, growth is pro-poor when it manifests itself in a reduction of inequalities. The second approach, known as absolute, considers growth to be pro-poor when it is accompanied by a reduction in the poverty rate in absolute terms. Each of these approaches has major shortcomings and some researchers have tried to propose alternative approaches (Fosu; 2015; Duclos and Verdier-Chouchane; 2011).

Today, the most important and undoubtedly the most difficult issue is the measurement of pro-poor growth. More concretely, how to identify inclusive growth? How many percentage points of growth must the poor receive in order to qualify as pro-poor? In recent years, practitioners have developed a multitude of measures to empirically determine the impact of growth on the poor[1]. However, despite the wide range of indices, important work is still needed on this issue of quantifying pro-poor growth. Indeed, there is currently no agreed measure of pro-poor growth. One of the main limitations of these indices is that they focus only on the monetary dimension of poverty. For this reason, recent years have seen the emergence of an alternative approach to quantifying the pro-poor character of growth. This second approach generally uses econometric models to establish causality between pro-poor growth and indicators of well-being.

This paper is a continuation of the latter approach. More precisely, the purpose of this paper is to use a new estimation technique that will allow us to conclude with more rigor on the question of the sign of this relationship.

[1] Examples include the growth incidence curve (CIC) of Ravallion and Chen (2003); the poverty growth curve of (Son; 2004); the pro-poor growth index of Kakwani et al. (2000); and so on.

Overall, although the econometric approach has the advantage of including the non-monetary dimension of poverty, it is subject to measurement errors problems and endogeneity bias. The latter are mainly explained by the inherent limitations of the data. Indeed, despite considerable improvements in recent years, data on pro-poor growth continue to pose serious problems for researchers. The high number of missing observations and the presence of outliers are deplored. Much of the data is of poor quality and is subject to significant measurement errors. Despite efforts to improve (or expand) the existing data, they remain unsatisfactory. Data on pro-poor growth differ between countries in terms of geographical coverage (national, urban or rural), statistical units (families, households or individuals) and definition of income (consumption expenditure, disposable income or gross income).[2]

One of the particularities of this paper is precisely to use a method for correcting these limitations. For this purpose, we determine, empirically, this relationship between poverty, growth and income inequalities in 24 developing countries by mobilizing a panel data over the period 2000-2007 and using the fixed effect Gini estimator. The latter, introduced by Ka and Mussard (2016), enables traditional hypotheses to be relaxed such as the linearity of the model[3]. Moreover, it allows controlling individual and temporal specifics effects and solving variables endogeneity bias[4] and measurement errors problems and provides more precise estimates of the effects of inequality and growth on poverty reduction in developing countries.

The rest of the paper is organized as follows: The next section presents the empirical specification and the data of this study. Section 3 describes the estimation method while section 4 the estimation results and section 5 concludes the paper.

2. Empirical specification and data

2.1 Empirical specification

In order to model pro-poor growth, we regress the first quantile (Q_1) on a set of variables such as growth rate (Txc), income inequality (Gini index), the logarithm of agricultural GDP per capita (log(GDPagr)) and the analysis includes control variables such as inflation rate (Inf), primary and secondary school enrolment rate (Eduprim and Edusec), public health expenditure (PH) and the share of public

[2]Moreover, in most empirical studies, the samples are composed of about sixty countries, among which developed countries are over-represented and poor countries, especially those in sub-Saharan Africa, almost non-existent, which leads to very heterogeneous samples and a drastic selection. Also, the temporal dimension of the sample is extremely limited relative to its cross-sectional dimension. As a result, the inter-individual variance is much higher than the inter-temporal variance (Marrero and Servén; 2018; Cogneau et al.; 2002; Ghura et al.; 2002). In addition, some authors (Barro; 2000; Ghura et al.; 2002) show the sensitivity of the results to the type of sample and the functional forms chosen.

[3]The estimators obtained by this approach are insensitive to the functional forms.

[4]Gini fixed effects Regression can be interpreted as an instrumental variable regression where the rank vectors of each regressor corresponds to the instrument matrix.

expenditure (SPE).

$$Q_{1ij} = \beta_0 + \beta_1 \text{Txc}_{ij} + \beta_2 \text{Gini}_{ij} + \beta_3 \text{Eduprim}_{ij} + \beta_4 \text{Edusecon}_{ij} + \beta_5 \text{Inf}_{ij} \\ + \beta_6 (1 - \text{Gini}_{ij}) * \text{Txc}_{ij} + \beta_7 \log(\text{GDPagr})_{ij} + \beta_8 \text{SPE}_{ij} + \beta_9 \log(\text{PH})_{ij} + \varepsilon_{ij} \quad (1)$$

The originality of this specification is multiple. First, unlike traditional approaches, this type of specification prevents us from defining a poverty line. Indeed, estimating a poverty line for all countries in our sample is a risky exercise: the information available is often not sufficient to define poverty lines rigorously based on the cost and basic needs technique. Also, most empirical work on pro-poor growth attempts to estimate the elasticity of the poverty rate. [5]

2.2 Data

To assess this relationship, we use an unbalanced panel data between 2000 and 2007 for 24 developing countries. Table B2 in Appendix B contains a variable description with their sources and summary description. Appendix B3 displays a list of the 24 developing countries included in this study.

The data are drawn from four sources. The first quantile and the Gini index come from the latest version of the World Income Inequality Database (WIID)[6]. Firstly, more recently, it allows to have many more observations (11 000 observations compared to 5314 from the first version of the WIID). Also, the procedure leading to the selection of data is much less constrained and thus allows more reliable cross-checking with other databases (the Luxembourg Income Study in particular). Another particularity of this last version is that it contains, in addition to the Gini index, the deciles, quantiles and percentiles P5 and P100. However, like the other databases on inequalities, the missing observations are extremely numerous.

The agricultural GDP per capita and growth rate variables come from the United States Department of Agriculture (USDA) database. These data were compiled by the Economic Research Service (ERS). It contains the GDP and growth rates of more than 194 countries and covers the period 1969 - 2015. Furthermore, all these observations are adjusted for inflation and have been reported in terms of the 2015 US dollar.

The secondary and primary education variables come from the Barro and Lee (2013) database. The other three analytical variables are taken from the World Development Indicators (WDI): the inflation rate, public health expenditure and the share of public expenditure.

[5] Obviously, we recognize that it is important to estimate this elasticity but unfortunately it is not sufficiently informative. Indeed, it does not tell us anything about the depth of poverty. Finally, this specification better responds to the question of whether the poor took benefit from the growth?

[6] The latest version of the WIID, released in May 2020, covers 200 countries (including historical entities), with over 11,000 data points in total.

3. Empirical Methodology or Estimation methods

We turn to the estimation of (1) by using different econometric approaches. The fixed effects ordinary least squares (OLS) estimators are very popular and convenient for empirical investigations, however, variables endogeneity bias, measurement errors problems and its sensitivity to model specification can drastically affect the estimates. In this setting, it is not obvious that OLS should be preferred to new estimation methods, such as Gini fixed effects estimator. The latter enables traditional hypotheses to be relaxed such as the linearity of the model. Moreover, it allows controlling individual and temporal specifics effects and solving variables endogeneity bias and measurement errors problems and provides more precise estimates of the effects of inequality and growth on poverty reduction in developing countries (Ka; 2016). For this reason, we display both sets of estimates (OLS and Gini), which help to assess the robustness of the results.

Formally, Gini regression consists of using the Gini mean difference (GMD) as a measure of dispersion:

$$GMD = \mathbb{E}\left|\mathbf{x}_i - \mathbf{x}_j\right| = 4\mathrm{Cov}(\mathbf{x}, F(\mathbf{x})),$$

where $F(\mathbf{x})$ stands for the c.d.f. of the random variable $\mathbf{x}$.

The latter was introduced in 1912 by Corrado Gini and since then several operators or coefficients have been deduced such as covariance in the Gini sense (co-Gini), correlation in the Gini sense (Gcorrelation), Gini analysis (ANOGI), Gini regression, Unit root test, heteroscedasticy test (Yitzhaki and Schechtman; 2013; Shelef; 2016; Charpentier et al.; 2019) etc.

Consider a model $\mathbf{y} = a + b\mathbf{x}$ with $\mathbf{x},\mathbf{y}$ some $N \times 1$ vectors. The semi-parametric Gini (simple) regression introduced by Olkin and Yitzhaki (1992), consists in averaging tangents b_{ij} (between observations i and j) with weights v_{ij}. Let the values of $\mathbf{x}$ be ranked by ascending order ($x_1 \leqslant \cdots \leqslant x_N$), then the semi-parametric Gini estimator of the slope coefficient is given by:

$$\hat{b}^G = \sum_{i<j} v_{ij} b_{ij}, \text{ with } v_{ij} = \frac{(x_i - x_j)}{\sum_{i<j}(x_i - x_j)} \text{ and } b_{ij} = \frac{(y_i - y_j)}{(x_i - x_j)} \; \forall i < j \; ; \; i = 1, \ldots, N.$$

The authors also demonstrate that if the weights v_{ij} are replaced by quadratic ones such as $w_{ij} = \frac{(x_i - x_j)^2}{\sum_{i<j}(x_i - x_j)^2}$, then the standard OLS estimator of the slope coefficient is obtained: $\hat{b}^{OLS} = \sum_{i<j} w_{ij} b_{ij}$. Since it depends on quadratic weights, the OLS slope coefficient is shown to be heavily sensitive to outliers.

The semi-parametric Gini regression may be defined according to the cogini operator, i.e. $\mathrm{cog}(\mathbf{y},\mathbf{x}) := \mathrm{cov}(\mathbf{y}, \mathbf{R}(\mathbf{x}))$ and $\mathrm{cog}(\mathbf{x},\mathbf{x}) := \mathrm{cov}(\mathbf{x}, \mathbf{R}(\mathbf{x}))$ where $\mathbf{R}(\mathbf{x})$ is the rank vector of $\mathbf{x}$ [7]:

$$\hat{b}^G = \frac{\mathrm{cov}(\mathbf{y},\mathbf{R}(\mathbf{x}))}{\mathrm{cov}(\mathbf{x},\mathbf{R}(\mathbf{x}))} = \frac{\mathrm{cog}(\mathbf{y},\mathbf{x})}{\mathrm{cog}(\mathbf{x},\mathbf{x})}, \text{ whereas } \hat{b}^{OLS} = \frac{\mathrm{cov}(\mathbf{y},\mathbf{x})}{\mathrm{cov}(\mathbf{x},\mathbf{x})}.$$

[7] The rank vector of $\mathbf{x}$ (of size $N \times 1$) is obtained by replacing the elements of x by their rank (the

The semi-parametric Gini multiple regression depends on the rank matrix of the regressors. Let $\mathbf{X}$ be the $N \times K$ matrix of the regressors and $\mathbf{R_x}$ its rank matrix, which contains in columns the rank vectors $\mathbf{R}(\mathbf{x}_k)$ of the regressors $\mathbf{x}_k$ for all $k = 1, \ldots, K$. The semi-parametric Gini multiple regression yields the following estimator (a $K \times 1$ vector):

$$\hat{\mathbf{b}}^G = (\mathbf{R}_\mathbf{x}'\mathbf{X})^{-1}\mathbf{R}_\mathbf{x}'\mathbf{y}$$

The semi-parametric Gini estimator is equivalent to that of instrumental variables regression in which the instruments are the rank vectors of each regressor (Durbin; 1954; Yitzhaki and Schechtman; 2004). Furthermore, the use of the Gini methodology enables one to see whether the conclusion reached suffers from deficiencies that originate from some of the hidden assumptions of the OLS [8] (Yitzhaki and Schechtman; 2013).

In line with the above literature, Ka and Mussard (2016) develop Gini estimators for panel data. They propose to decompose the variability of the moment matrices into within- and between-group Gini variabilities in order to deduce a fixed effects Gini regression for panel data. They show that the within-group Gini estimator derived from this decomposition is a Gini estimator. It is also an U-statistics[9], consequently, it is asymptotically normal. Thus, to derive more general conclusions about the relationship between taxes and income inequality in developing countries, we use OLS and Gini estimators.

4. Estimation results and interpretation

The Tables B4 and B5 in Appendix report the results of our estimations with the standard fixed effect and the Gini fixed effect, respectively.

In Table B4, we have done four estimations. The first and second columns of this Table displays the results of regression using a within- and between-group estimator, and the third and last column show the results using OLS and first difference estimators (FD). In the OLS regressions, we observe that Growth rate (at the 5% level) have positive and significant relationship with reduction poverty whereas on the other hand, inflation rate and Gini index (at the 5% level) have a significant negative relationship with poverty reduction. Surprisingly, log(GDPagr) has an unsignificant positive relationship with poverty reduction. This result is not in line with

smallest value of $\mathbf{x}$ being 1 and the highest being N). It is worth mentioning that for ties in the regressors, we have to estimate the values of the rank vector as mid-points. The procedure is similar to the case of weighted samples, see (Yitzhaki and Schechtman; 2013).

[8] The Gini semi-parametric approach has the advantage of relying on a few assumptions and no linearity hypothesis is needed.

[9] Yitzhaki and Schechtman (2013) show that all the estimators used in Gini regressions are U-statistics and an easier way to estimate its variance is the jackknife:

$$Var(\mathbb{U}) = \frac{N-1}{N}\sum_{i=1}^{N}\left[\mathbb{U}_{-i} - \frac{1}{N}\sum_{i=1}^{N}\mathbb{U}_{-i}\right]^2,$$

where $\mathbb{U}_{-i}$ is the estimator based on a sample of size N, without the ith observation.

the classical theory, which predicts that agriculture plays an important role in poverty reduction (Christiaensen and Demery; 2007; Mosley and Suleiman; 2007). It supports other conflicting literature such as Christiaensen et al. (2006) and Machethe (2004). It is probably due to a possible endogeneity of the explanatory variables or, more generally, to a misspecification of the functional form chosen.

The sensitivity of standard fixed effect estimators to endogeneity problem, measurement errors and functional forms means that we will prefer to use the within-group Gini regression estimator. The latter is more robust than the within estimator based on OLS (Ka and Mussard; 2016).

Next, Table B5 presents the regression results with the Gini methodology. Similarly, the first and second columns report the result of regression using the within- and between-group Gini estimators, and the third and last columns report the result using the Gini and first difference Gini estimator. These results lead to quite significantly different assessments of the impact of variations in growth rates and inequalities. There is even a sign opposition for the estimated parameters of the log(GDPagr). These discrepancies are mainly due to the sensitivity of the usual estimator to the model specification, but also to large measurement errors, outliers[10] (Table B1 in Appendix) and endogeneity.

Overall, we note across Table B5 that the coefficient of the inflation rate, growth rate, Gini index, and log(GDPagr) are significant at the 1% level. This result is compatible with the political economy approach, according to which increasing inequality leads to greater social pressure towards distribution policies. These policies generate distortions that harm capital accumulation, and then poverty reduction (Delbianco and Dabús; 2009; Christiaensen and Demery; 2007). Also, according to this standard literature, inflation has a negative impact on the lower quantiles. It erodes the purchasing power of the poor. In reality, inflation can be seen as a regressive tax that affects the poorest more (Chani et al.; 2011).

Furthermore, in recent decades, there has been an abundance of literature on the importance of access to education in the fight against poverty. Indeed, education plays an essential role in finding employment and achieving personal autonomy. However, although the secondary school enrolment rate (Edusec) plays a positive role in poverty reduction, the analysis reveals that the primary school enrolment rate (Eduprim) does not have a significant influence on the lower quantiles. The main explanation for this result is that stimulating access to primary education is effective in reducing poverty only if the populations concerned can continue their education in order to benefit fully from the high marginal returns associated with long-term training.

[10]Probably the most popular tools for detecting outliers are cook's distance and DFBETAS. The latter indicates the effect that deleting each observation has on the estimates for the regression coefficients. Values larger than $2/sqrt(N)$ (with N=123, |DFBETAS| > 0.180) in absolute value are considered an outlier. Cook's distance indicates the effect that deleting each observation has on the predicted values of the model. Values larger than $4/N$ (with N=123, Cook's distance > 0.032) are considered highly influential.

It is interesting to note that, contrary to OLS, the results obtained in the different dimensions of variability in the Gini sense lead us to relatively close estimates. Indeed, Gini, within- and between-group estimators lead to assessments of the impact of growth and inequality on fairly identical quantiles, which is an additional argument in favor of Gini regression. In reality, the relative proximity of the estimators resulting from the different centering methods generally constitutes a validation of the model specification. Thus, Gini regression should not be reduced to a simple technique for dealing with outliers. As we have seen previously, this regression technique has several other advantages: it is unbiased, convergent, less sensitive to the model specification and to measurement errors.

5. Conclusion

The aim of this paper is to examine the impact of growth and inequality on poverty reduction in developing countries by mobilizing unbalanced panel data over the period 2000-2007 and using the Gini methodology which accounts for endogeneity issues and measurement errors. It is the first paper to use this methodology in the empirical literature on pro-poor growth. Our main results suggest that income inequality and inflation negatively impact poverty reduction and there is a positive and robust relationship between growth and the logarithm of agricultural GDP per capital on poverty. Also, the results indicate that Gini methodology addressing the econometric shortcomings of OLS regression analysis (endogeneity bias, outliers, and measurement errors) may yield more precise results.

References

Anand, S. and Kanbur, S. R. (1993). Inequality and development a critique, *Journal of Development economics* **41**(1): 19–43.

Barro, R. J. (2000). Inequality and growth in a panel of countries, *Journal of economic growth* **5**(1): 5–32.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010, *Journal of development economics* **104**: 184–198.

Bourguignon, F. (2003). The growth elasticity of poverty reduction: explaining heterogeneity across countries and time periods, *Inequality and growth: Theory and policy implications* pp. 3–26.

Chani, D., Irfan, M., Pervaiz, Z., Jan, S. A., Ali, A. and Chaudhary, A. R. (2011). Poverty, inflation and economic growth: empirical evidence from pakistan, *World Applied Sciences Journal* **14**(7): 1058–1063.

Charpentier, A., Ka, N., Mussard, S. and Ndiaye, O. H. (2019). Gini regressions and heteroskedasticity, *Econometrics* **7**(1): 4.

Christiaensen, L. and Demery, L. (2007). *Down to earth: agriculture and poverty reduction in Africa*, The World Bank.

Christiaensen, L., Demery, L. and Kuhl, J. (2006). *The role of agriculture in poverty reduction an empirical perspective*, The World Bank.

Cogneau, D., Guénard, C. et al. (2002). Les inégalités et la croissance: une relation introuvable, *Document de travail DT* **3**.

Delbianco, F. and Dabús, C. (2009). María angeles caraballo pou (2016): Growth, inequality and corruption: Evidence from developing countries, *Economics Bulletin* **36**(3): 1811–1820.

Duclos, J.-Y. and Verdier-Chouchane, A. (2011). Analyzing pro-poor growth in southern africa: Lessons from mauritius and south africa, *African Development Review* **23**(2): 121–146.

Durbin, J. (1954). Errors in variables, *Revue de l'institut International de Statistique* pp. 23–32.

Fosu, A. K. (2015). Growth, inequality and poverty in sub-saharan africa: recent progress in a global context, *Oxford Development Studies* **43**(1): 44–59.

Ghura, D., Leite, C. and Tsangarides, C. G. (2002). *Is growth enough? Macroeconomic policy and poverty reduction*, Vol. 2, International Monetary Fund.

Ka, N. (2016). *Modélisation de la croissance pro-pauvre*, PhD thesis.

Ka, N. and Mussard, S. (2016). l1 regressions: Gini estimators for fixed effects panel data, *Journal of Applied Statistics* **43**(8): 1436–1446.

Kai, H., Hamori, S. et al. (2009). Globalization, financial depth and inequality in sub-saharan africa, *Economics Bulletin* **29**(3): 2025–2037.

Kakwani, N., Pernia, E. M. et al. (2000). What is pro-poor growth?, *Asian development review* **18**(1): 1–16.

Klasen, S. (2005). Economic growth and poverty reduction: measurement and policy issues.

Machethe, C. L. (2004). Agriculture and poverty in south africa: Can agriculture reduce poverty, *Paper presented at the Overcoming Underdevelopment Conference held in Pretoria*, Vol. 28, Citeseer, p. 29.

Marrero, G. A. and Servén, L. (2018). *Growth, Inequality, and Poverty: A Robust Relationship?*, The World Bank.

Mosley, P. and Suleiman, A. (2007). Aid, agriculture and poverty in developing countries, *Review of Development Economics* **11**(1): 139–158.

Olkin, I. and Yitzhaki, S. (1992). Gini regression analysis, *International Statistical Review/Revue Internationale de Statistique* pp. 185–196.

Ravallion, M. and Chen, S. (2003). Measuring pro-poor growth, *Economics letters* **78**(1): 93–99.

Shelef, A. (2016). A gini-based unit root test, *Computational Statistics & Data Analysis* **100**: 763–772.

Son, H. H. (2004). A note on pro-poor growth, *Economics Letters* **82**(3): 307–314.

Yitzhaki, S. and Schechtman, E. (2004). The gini instrumental variable, or'the double iv estimator', *Metron* **LXII**(3): 287–313.

Yitzhaki, S. and Schechtman, E. (2013). *The Gini methodology: A primer on a statistical methodology*, Vol. 272, Springer Science & Business Media.

A. Appendix : Gini R-squared and outlier tests

A.1 Gini R-squared

Following Olkin and Yitzhaki (1992), the Gini R-squared (GR^2 hereafter) is given by

$$GR^2 = 1 - \left(\frac{\mathrm{Cov}(\mathbf{e},\mathbf{R}(\mathbf{e}))}{\mathrm{Cov}(\mathbf{y},\mathbf{R}(\mathbf{y}))}\right)^2,$$

where $\mathbf{R}(\mathbf{e})$ is the rank vector of $\mathbf{e} = \mathbf{y}_t - \hat{\mathbf{y}}_t$ and $\mathbf{R}(\mathbf{y})$ the rank vector of $\mathbf{y} = \mathbf{y}_t - \bar{\mathbf{y}}_t$.

B. Appendix : Summary Statistics and Estimations

B.1 Outlier tests

Table B1
Outlier tests

Obs	Cook's distance	DFBETAS Eduprim	Edusecon	inf	HE	Gini	(1)	Txc
25	2.63***	−5.094***	0.303***	−0.201***	0.732***	−0.415***	1.189***	1.023***
35	3.130***	−2.309***	0.504***	-0.306	1.031***	−0.298***	1.952***	2.044***
62	1.012***	−1.022***	0.652***	−0.521***	2.675***	−0.865***	1.523***	2.157***
82	1.012***	−2.042***	0.320***	−0.234***	1.024***	−1.04***	1.620***	1.956***
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
112	1.23***	1.101***	0.712***	0.502***	2.018***	0.810***	1.012***	1.247***

(1) represents Log(GDPagr). Note also *** indicates values larger than cuttof value of cook's distance and DFBETAS.

B.2 Definition, Summary Statistics and Country list

Table B2
Source and Descriptive statistics

Variable	Definition	Source	Mean	Std. Dev.
Gini index	Gini index	WIID	46.759	8.141
Q1	First Quantile	WIID	4.576	1.905
Edusec	Secondary school enrolment rate	(2)	58.236	26.809
Eduprim	Primary school enrolment rate	(2)	101.946	19.35
inf	Inflation, GDP deflator (annual %)	WDI	10.068	11.369
HEP	Public health expenditure (% of GDP)	WDI	2.662	1.202
Growth	Growth rate	USDA	4.612	5.019
Log(GDPagr)	log of agricultural GDP per capita (constant 2015 USD)	USDA	2.963	4.901

(2) represents Barro and Lee (2013)

Table B3
Country list

Country	Years included	Country	Years included
Albania	2000,2003, 2004, 2005	Jamaica	2003-2007
Armenia	2000-2007	Madagascar	2003,2005,2006,2007
Bolivia	2000-2004,2006,2007	Malawi	2003,2005,2006,2007
Burkina Faso	2003,2004,2006,2007	Mauritania	2003,2005,2006,2007
Cameroon	2003,2005,2006,2007	Morocco	2003-2007
Dominican Republic	2000, 2003,2006	Mozambique	2003,2005,2006,2007
Ecuador	2003-2007	Nigeria	2001-2007
El Salvador	2000, 2001, 2003-2006	Paraguay	2003,2004,2005,2007
Ethiopia	2000-2007	Senegal	2004,2006,2007
Ghana	2000-2007	Uganda	2000-2006
Guatemala	2000, 2001, 2004-2007	Uruguay	2003,2004,2005,2007
Honduras	2000-2007	Zambia	2003,2007

B.3 Estimations

Table B4
OLS Esimates

Estimates → $\beta =$	OLS Esimates Within	Between	OLS	First-difference (FD)
Eduprim	-0.075 (0.081)	0.050 (0.051)	0.037 (0.042)	0.011 (0.012)
Edusecon	0.070^{*} (0.042)	0.048^{*} (0.028)	0.021^{*} (0.012)	0.15^{*} (0.091)
Inflation rate	-0.034^{**} (0.017)	-0.056^{**} (0.028)	-0.016^{**} (0.008)	0.001^{**} (0.001)
Public health expenditure	0.013 (0.016)	0.035 (0.073)	0.011 (0.027)	0.026 (0.029)
Gini index	-0.117^{**} (0.004)	-0.065^{**} (0.029)	-0.224^{**} (0.092)	-0.162^{**} (0.075)
Log(GDPagr)	-0.041 (0.042)	0.182 (0.153)	0.086^{**} (0.051)	-0.063^{**} (0.037)
Growth rate	0.145^{**} (0.057)	0.181^{**} (0.071)	0.057^{**} (0.023)	0.23^{***} (0.027)
(1-Gini)*Txc	0.051^{***} (0.003)	0.056^{***} (0.005)	0.070^{***} (0.019)	0.065^{***} (0.053)
Share of public expenditure	-0.011 (0.05)	0.099 (0.086)	0.055 (0.081)	0.087 (0.078)
R^2	0.627	0.887	0.889	0.721

Note: *, ** and ***: 10. 5 and 1% significant respectively, Standard errors are in parenthesis.

Table B5
Gini methodology Estimates

Estimates →	Gini methodology Estimates			
$\beta =$	Within Gini	between Gini	Gini regression	FD Gini
Eduprim	0.011 (0.07)	0.012 (0.051)	0.011 (0.085)	0.009 (0.032)
Edusecon	0.087^{**} (0.033)	0.094^{**} (0.045)	0.091^{**} (0.038)	0.065^{**} (0.029)
Inflation rate	-0.089^{***} (0.001)	-0.099^{**} (0.003)	-0.097^{***} (0.0012)	0.101^{***} (0.001)
Public health expenditure	-0.011 (0.046)	0.013 (0.089)	0.120 (0.911)	0.009 (0.120)
Gini index	-0.116^{***} (0.009)	-0.13^{***} (0.011)	-0.141^{***} (0.012)	-0.15^{***} (0.025)
Log(GDPagr)	0.097^{***} (0.005)	0.119^{***} (0.009)	0.117^{***} (0.006)	0.179^{**} (0.022)
Growth rate	0.110^{***} (0.0070)	0.095^{***} (0.012)	0.086^{***} (0.017)	0.120^{***} (0.019)
(1-Gini)*Txc	0.039^{***} (0.001)	0.045^{***} (0.002)	0.042^{***} (0.001)	0.021^{**} (0.004)
Share of public expenditure	0.013 (0.055)	0.016 (0.033)	0.014 (0.057)	0.026 (0.045)
GR^2	0.582	0.751	0.792	0.761

Note: *, ** and ***: 10. 5 and 1% significant respectively, Standard errors are in parenthesis. FD represents first-difference.