Journal of African Economies, 2023, **32**, ii296–ii319
https://doi.org/10.1093/jae/ejac053
Supplement Paper

OXFORD

Income Inequality and Redistribution in Sub-Saharan Africa

Miguel Niño-Zarazúa[a,b,*], Francesca Scaturro[c], Vanesa Jordá[d] and Finn Tarp[e]

[a]Department of Economics, SOAS University of London, London, United Kingdom
[b]UNU-WIDER, Helsinki, Finland
[c]Department of Economic and Social Sciences, Università Politecnica delle Marche, Ancona, Italy
[d]Department of Economics, University of Cantabria, Santander, Spain
[e]Department of Economics, University of Copenhagen, Copenhagen, Denmark

*Corresponding author: Miguel Niño-Zarazúa. E-mail: mn39@soas.ac.uk

Abstract

A strand of the political economy literature emphasizes the effect of income inequality on growth and poverty, which materialises through redistribution. The theoretical expectation postulated by standard economic theory is that high inequality would lead to higher redistribution via the collective action of the median voter. The empirical literature testing the *median voter theorem* has been conducted in the context of industrialised economies. In this article, we examine the median voter hypothesis with specific reference to SSA, a region characterised by high levels of income inequality and limited redistribution. We adopt an instrumental variable approach to unpack the determinants and plausible mechanisms underpinning this relationship. In the article, we account for the effect of omitted top income earners in income inequality estimates, given their weight in the shape of the income distribution and their influence in redistributive policies. Overall, we find a positive relationship between inequality and redistribution in SSA, especially among middle-income countries. Further examination reveals that the abundance of natural resource rents seems to be the driving force affecting tax policy choices, which in turn exacerbates income inequality and undermines progressive redistribution. Thus, in assessing the relationship between inequality and redistribution, our results do not provide strong evidence to support the propositions of the median voter theorem, but instead, they call for alternative interpretations that seem to align more closely to the existence of multiple steady states.

Keywords: SSA, taxation, redistribution, inequality
JEL classification: D63, D72, E62, H20, H39

1. Introduction

High levels of income inequality in many parts of the developing world have drawn the attention of scholars to investigate their drivers and consequences, and the extent to which the median voter and poorer members of society influence governments' redistributive decisions (McCarty and Pontusson, 2011). One of the main concerns about high levels of income inequality is the possible negative effects on economic growth and aggregate welfare.

Indeed, a long-standing debate exists in the economics literature about the impact of income inequality on economic and social development (Adelman and Robinson, 1989). The pioneering works by Kuznets (1955) and Lewis (1954) provide a theoretical analysis of the underlying mechanisms in the relationship between inequality and economic development, focusing on the sectoral composition of the economy. Specifically, inequality is expected to increase with the shift from a low-income agrarian economy to a high-income modern industrialised economy.

As for the possible influence of inequality on growth, a positive association, that is inequality as growth enhancing, has been envisaged based on three main arguments (Aghion *et al*., 1999). First, the rich have higher marginal propensity to save, which translates into higher aggregate savings and growth. Second, the existence of investment indivisibilities in the presence of imperfect capital markets requires some concentration of wealth to finance certain productive activities. Third, the existence of incentives would foster the production of output when the latter depends on effort.

By assuming a different perspective, some studies point out the detrimental effects of inequality on growth. Among others, Galor and Zeira (1993) and Banerjee and Newman (1993) looked at the role of credit market imperfections. Specifically, they highlight how credit constraints reduce the ability of the poor to invest in education, which in turn impact occupational choices, labour productivity and create income gaps that ultimately hampers aggregate output.

A much smaller strand of the literature that emphasises the effects of income inequality takes a political economy perspective. The seminal works by Alesina and Rodrik (1994) and Persson and Tabellini (1994) highlight a negative effect of inequality on growth, which materialises through redistributive policies. In contrast, Li and Zou (1998) come to a different conclusion, according to which inequality have a positive effect on growth.[1]

Taxing the wealthy would have two effects on growth: one would reduce the net return on production factors, such as capital and skilled labour, thus affecting negatively growth. Another would increase transfers to the poor and finance public services such as infrastructure and education that would stimulate growth. Since redistribution decisions are endogenous to inequality, past inequality would influence redistribution and consequently, future economic growth.

While the theoretical predictions from this strand of literature are certainly relevant for developing countries, the empirical evidence testing these dynamics remains largely ambiguous. In this article, we contribute to fill this gap by examining the relationship between income inequality and redistributive decisions, particularly in the context of sub-Saharan Africa (SSA), a region characterised by high levels of income inequality and limited redistribution. We adopt an instrumental variable approach to unpack the determinants and likely mechanisms underpinning the association between income inequality and redistribution. Given the role of elites highlighted by the literature as influencing redistributive decisions, we follow Jorda and Niño-Zarazúa (2019) to account for the effect of omitted top incomes in the estimation of income inequality due to existing data constraints in household surveys.

Overall, we find strong evidence of a negative effect of inequality on total government revenues, our proxy for redistribution. The results are consistent for most country income groups, and across model specifications, econometric methods and inequality measures, with the only exception of SSA, which differs from the rest of the global sample by showing a positive effect of inequality on redistribution. Interestingly, accounting for the effect of omitted top incomes (those at the top 99 percentile of the income distribution) on the level of income inequality has a qualitatively negligible effect on redistribution. This seems to reflect not only a limited revenue mobilisation capacity via direct taxes of SSA countries, but also

[1] See Ostry *et al*. (2014) for a formal discussion on the relationship between inequality, redistribution and growth.

the likely strength of elite cohesion and their connectedness with political regimes, which in the presence of natural resources rents, undermine the feasibility of progressive taxation. Thus, our results do not seem to provide strong evidence to support the propositions of the median voter theorem, but instead, they hint at alternative propositions that underpin the causal relationship between income inequality and redistribution in low- and lower–middle-income countries.

The remainder of the article is organised as follows: the next section reviews the literature with specific reference to the redistribution hypothesis, particularly in the context of SSA. Next, we present the empirical strategy and the model specification while highlighting the relationship between inequality and redistribution. Next, we describe the data sources and key variables used in the empirical analysis before moving on to the results. The next section presents a series of robustness checks while the final section concludes.

2. Inequality and redistribution

Within the political economy literature, there is an emphasis on the role of the median voter in influencing redistribution decisions, particularly in the context of high levels of inequality (Meltzer and Richard, 1981; Alesina and Rodrik, 1994; Persson and Tabellini, 1994). The theoretical expectation is that in contexts of a competitive electoral systems, high inequality would lead to higher redistribution via the collective action of the median voter.

Most of the empirical literature testing the role of inequality in influencing redistribution via the median voter has been conducted in the context of advanced economies, most of them with long standing liberal democracies, providing mixed results. Studies that support a positive association between inequality and redistribution (Shelton, 2007; Boustan *et al.*, 2013) differ in terms of sample, timeframe, proxies for both inequality and redistribution, and estimation strategies, making the comparison of findings intricate.[2] To illustrate, redistribution has been measured by the difference in the share of the bottom quantiles of the income distribution when disposable income is considered in relation to factor income (Milanovic, 2000) or, by the change in the Gini coefficient from gross market income to disposable income (Lupu and Pontusson, 2011; Scervini, 2012; Luebker, 2014). Further analyses have been conducted using social spending or tax revenues as proxy measures for redistribution (Schwabish *et al.*, 2006).

Several studies that examine the association between inequality and redistribution, do not find any significant result (see De Mello and Tiongson (2003) for a review), while others report a non-positive (Lindert, 1996) or a non-linear (De Mello and Tiongson, 2003) relationship. It should be noted that some of the conditions necessary for the median voter theorem to apply hardly hold for developing countries whose political institutions and electoral systems differ in significant ways from those outlined by the median voter model. Even among liberal and consolidated democracies, it is not always the case that countries with high levels of income inequality redistribute more.

In light of the inconclusive evidence from the literature, it is pertinent to consider alternative interpretations of the relationship between inequality and redistribution. From a ‘social contract paradigm’, Bénabou (2000) considers a non-linear relationship between inequality and redistribution with multiple steady states: one with high inequality and low redistribution and the other with low inequality and high redistribution. The rationale of the model is that in the presence of fragmented credit and insurance markets, redistribution will have an insurance and a productivity enhancing effect, even if it leads to a reduced economic efficiency. In contexts with low levels of inequality, redistributive policies will have the popular support. However, as inequality increases, the influence of rich population

2 Scervini (2012) reviews some of the most influential studies of the early reference literature.

will become sufficiently strong to oppose further redistribution. In the end, when inequality reaches high levels, the share of poor population will become large enough to demand an increasing redistribution, even if it is inefficient.

Similarly, the work by Moene and Wallerstein (2001) predicts a negative relationship between inequality and redistribution. In this case, however, behind such a negative association, there is the assumption that social spending is not only a way to redistribute income but also to provide some forms of insurance.

More recently, other interpretations of the mechanisms underlying the median voter hypothesis have been proposed. In particular, the rational utility maximisation paradigm driving the median voter's choice in the traditional approach has been revised on the basis of arguments from behavioural economics emphasising the role of individual incentives and normative value judgements in shaping preferences about redistribution (Luebker, 2014; Ahrens, 2019; Bussolo *et al.*, 2019b). In addition, taking advantage of the substantial improvement in the quality of data, empirical analyses on the political economy of redistribution have been increasing. The social contract paradigm has been more recently tested also with reference to developing countries. Prominent analyses are those by Breceda *et al.* (2009) for Latina America, Birdsall and Haggard (2002) for East Asia, and Zoellick (2011) for Middle East and North Africa.[3]

3. Empirical strategy

Empirical analyses of the relationship between inequality and redistribution remains ambiguous partly due to two important constraints: first, data has been a major limitation, especially for cross-country analysis. Second, some of the underlying assumptions of the median voter theorem are difficult to test in developing country contexts, partly because social and political institutions may differ from the assumptions imposed by the theorem. In order to test this hypothesis empirically, we present in the next section the baseline model.

3.1. Model specification

In order to assess the effect of inequality on redistribution, we estimate the following model:

$$R_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \upsilon_t + c_{it} \tag{1}$$

where the subscripts i and t denote country and period, respectively, R_{it} is a proxy for redistribution, β_0 is the constant, I_{it} is an index of income inequality, X is the matrix of the control variables, υ_t is a vector of period dummies capturing common time trends and c_{it} is the error term.

Inequality (I_{it}) is our key variable of interest. Specifically, we want to assess whether, and the extent to which, income inequality affects redistributive decisions. It should be noted here that inequality is likely to be endogenous in Equation (1) due to several reasons. First, the presence of omitted variables influencing both inequality and redistribution. Second, the possibility of measurement error due to the absence of top income earners in household surveys. Finally, simultaneity bias may emerge since the level of inequality is likely to influence redistribution as much as redistribution is likely to influence the level of inequality. In such cases, the assumption of exogeneity would not hold and we would need to find a valid instrument for inequality to make our estimates consistent. Consequently, we extend

[3] It should be acknowledged that the debate about social contract is still open and the related literature is still evolving, for both less developed and advanced countries. Among the most recent contributions, see Bussolo *et al.* (2019a).

Equation (1) into a system of equations, by modelling inequality as follows:

$$I_{it} = \delta_0 + \delta_1 Z_{it} + \delta_2 X_{it} + \upsilon_t + u_{it} \tag{2}$$

where Z_{it} is exogenous with respect to Equation (1), but partially correlated with inequality in Equation (2), i.e., Cov(Z_{it}, c_{it}) = 0 and δ_1 /= 0.[4] The variables considered as instruments for inequality in this analysis are described in Section 3.2.3. In addition to inequality, we control for other factors that influence redistributive decisions, following the reference literature (see Dioda (2012), Drummond *et al.* (2012), Sen Gupta (2007)). First, we consider some structural economic factors. As proxies for the level of economic development, we use both per capita income (yPPP) as well as the share of value added originating from agriculture (agric), the latter variable providing also information about the sectoral composition of output. Per capita income is expected to be positively correlated with government tax revenues–our proxy for redistribution–since the demand for goods and services provided by governments is expected to increase with income. In addition, economic development usually goes along with greater governments' capacity to levy and collect taxes (Dioda, 2012). In contrast, a high share of agriculture over national output denotes a less diversified and developed economy, which in turn negatively impact government revenues. Moreover, when characterized by subsistence farming and mainly driven by dispersed small-scale producers, the primary sector may also be difficult to tax (Sen Gupta, 2007).

We also include in the model an indicator that measures the trade openness of countries (trade), since the share of import and export over GDP is expected to influence the revenue performance of an economy and the size of the government, although the direction of its association with tax revenues remains ambiguous in the literature. On the one hand, taxes on imports and exports are relatively easy to collect because the monitoring of the entry and exit of goods into and from a country is generally straightforward, thus leading to a positive association with tax revenues. On the other hand, trade liberalization and trade agreements usually involve cuts in international tax rates which, in the absence of appropriate domestic tax reforms can result in a consequential fall in government revenues (Khattry and Rao, 2002; Gnangnon and Brun, 2019).

Furthermore, in order to control for the influence of the overall economic cycle, we include the unemployment rate (unempl). In principle, tax revenues are expected to rise during booms while falling during recessions. As a consequence, the correlation between tax revenues and unemployment would be expected to be negative, although the country-specific revenue composition and the procyclicality of fiscal policies characteristic of many developing countries may influence and even reverse the expected pattern of this relationship (Talvi and Vegh, 2005; Alesina *et al.*, 2008).

Second, we consider some socio-demographic factors influencing tax revenues. In particular, we control for the dependency ratio of countries (depratio), defined as the share of population younger than 15 or older than 64 to the working-age population (aged 15–64), as well as for female participation to the labour force (femlabpart). Both variables are expected to be positively associated with revenue collection, although not unambiguously (Dioda, 2012). Countries characterized by a high or rapidly growing proportion of their elderly population face the pressure to create or expand their pension systems, a goal which can be favourably approached through increasing revenues. In contrast, countries with a large proportion of children face limited productive capacity that generates tax revenues. Female labour force participation is expected to be positively correlated with tax revenue as a higher share of women employed in the labour market enlarges the tax base (Dioda, 2012).

[4] Following Andrews *et al.* (2019), we refer to equation (1) as the *structural form* equation and to equation (2) as the *first stage* equation.

We also control for population density (popdens), since it is expected to lower the administrative costs of tax collection and evasion controls. Finally, we consider ethnic tensions (ethnt), in order to assess whether ethnicity may affect the mobilization of collective resources and the provision of public goods (Alesina *et al.*, 1999). The literature has widely highlighted the influence of ethnic composition on countries' economic performance (Alesina and La Ferrara, 2005; Habyarimana *et al.*, 2007). Moreover, specific attention has also been devoted in examining the influence of ethnicity on the government effectiveness, with some studies arguing that individuals in diverse communities are less willing to contribute to the public good (Kimenyi, 2006; Lindqvist and Östling, 2013), while others find an ethnic diversity dividend (Gisselquist *et al.*, 2016). Nonetheless, ethnic fractionalization could lead to lower tax revenues, especially in countries characterized by an important colonial history, which might have resulted in fragmented policies and weaker national identities (Besley and Persson, 2014).

Third, we consider a set of institutional factors in the realm of the political system that may exert some influence on revenue collection (Bird *et al.*, 2014). Specifically, we include proxy indicators for government stability (govstab); i.e., the ability of governments to carry out their declared programmes and policies, internal conflict (intconfl), i.e., the political violence in the country and its actual or potential impact on governance, and corruption (corrup) within the political system. Overall, we expect higher institutional quality and political stability to positively influence revenues collection, while more corruption to be negatively associated with tax revenues (Botlhole *et al.*, 2012). In the next Section, we describe the main indicators used in the empirical analysis, and the data sources.

3.2. Data and variables

3.2.1. Revenues

We estimate model (1) by using total government revenue as share of GDP as our dependent variable. Total government revenue captures the level of fiscal resources available to governments and it is a valid approximation for a country's redistributive capacity. In fact, the ability to collect taxes is central to a country's capacity to finance social services such as health and education, critical infrastructure and other public goods (Akitoby *et al.*, 2019). Moreover, the correlation between redistribution and revenues has been widely documented (see Ostry *et al.* (2014)).

Given the international comparative perspective of the present analysis, we resort to UNU-WIDER's Government Revenue Dataset (GRD), which provides cross-national information on governments' revenue collection capacity. Specifically, we use the series of revenues exclusive of social contributions.[5] This choice is motivated based on the problems of completeness and comparability for social contribution figures, particularly for developing countries.

3.2.2. Inequality

We estimate the reference model (1) by using the Gini coefficient as our preferred measure of income inequality. The Gini index for each country and reference year were estimated using data on income shares from UNU-WIDER's World Income inequality Database (WIID), which contains repeated cross-country information on Gini indices and income (or consumption) shares for 189 countries.[6] The WIID is the most reliable and comprehensive database of worldwide distributional data currently available.[7]

Whenever we had missing information for every reference country-year data point, we opted to include observations within a maximum of the previous or next five years of each

[5] Revenues data used for the analysis are also exclusive of grants.
[6] The WIID database is available on the following link: https://www.wider.unu.edu/database/wiid.
[7] For a review of the data coverage and the main statistical features of the WIID, see Jenkins (2015).

data point, while giving preference to the closest observations. In addition, we adopted the conceptual base of the Canberra Group to minimize the problems that may arise from informational differences in the WIID in terms of unit of analysis, equivalence scale, the quality of the data and the welfare concept.[8]

In order to keep the global coverage as high as possible, we included consumption-based quintile data, in addition to income-based data, which is our preferred welfare concept. We note that mixing consumption and income data could lead to misleading results because both variables present different distributional patterns, being consumption typically characterized by lower inequality. Therefore, we adopt a harmonization procedure that consists of comparing the average income shares with those of consumption, for the available country-year observations that had both income and consumption data available for the same year. Then, we grouped countries into world regions and computed an average index of income relative to consumption, following Jorda and Niño-Zarazúa (2019). This procedure is similar, although not strictly identical to the ones adopted by Niño-Zarazúa *et al.* (2017) and Deininger and Squire (1996), with the key distinctive feature being that in the present study, we account for the difference in the income-consumption relationship at the regional, not global, levels.

An important potential source of bias in the empirical literature comes from the omission of top income earners in household surveys, from which inequality measures such as the Gini index are generated. The size of the national income pie in the hands of the richest can change not only the shape of the income distribution and the level of income inequality, but also governments' incentives and preferences for redistribution. Previous studies have used administrative records on personal income tax returns to adjust the upper tail of the income distribution coming from household surveys (Leigh, 2007; Atkinson *et al.*, 2011; Alvaredo *et al.*, 2013, 2017; Piketty and Saez, 2013). Tax records, however, are only available for a very small number of countries, and mostly for a relative short time window.

In order to overcome the limitations in the existing literature, we follow Jorda and Niño-Zarazúa (2019), and apply a parametric model, based on the so-called generalized beta distribution of the second kind (GB2) that helps us estimate the size of the bias–or truncation points in the Lorenz curves–arising from the omission of top incomes in the estimation of income inequality measures. We mitigate this bias by adjusting the income distribution after setting the truncation points at the $t = 0.99$, 0.9925, 0.995 and 0.9975 percentile levels. We then estimate the reference model (1) based on both the unadjusted Gini index and the Gini adjusted by top incomes, following the truncation points described above.

3.2.3. Instrumental variables

In order to control for the simultaneity bias problem in the relationship between inequality and redistribution, we experiment with three instrumental variables that have been used in previous studies. The first instrument captures countries' agricultural endowments. Following Easterly (2007), we consider the share of land used to produce wheat, relative to the share of land used for sugarcane production (wheatsugar). The rationale behind this instrument is motivated by Sokoloff and Engerman's (2000) hypothesis that the abundance of land for specific modes of agricultural production in former colonies set a pattern of structural inequality that continues to influence inequality levels in many developing countries, but it not expected to exert a direct influence on redistribution. We compute this

[8] More specifically, we focus on individuals rather than households, as the preferred unit of analysis. We also opt for income per capita rather than adult equivalent adjustments. In addition, we give preference to observations from nationally representative surveys, which are deemed to be of the highest quality. Finally, our preference is to use income over consumption as the welfare concept in the analysis.

instrumental variable as follows:

$$\text{wheatsugar} = \ln\left(\frac{1 + \text{wheat agril}}{1 + \text{sugarcane agril}}\right) \tag{3}$$

where *wheat agril* is the share of land used to grow wheat over total arable land while *sugarcane agril* is the share of land used to grow sugarcane over total arable land. We use lagged values of this indicator as instrument to current inequality. We expect a higher incidence of land for growing wheat to be associated with lower inequality. In fact, as pointed out by Easterly (2007), sugarcane was a labour-intensive crop compared to wheat, and its production proved to be profitable only in the presence of economies of scale obtained in large plantations. These features led nations with relative abundance in land suitable for sugarcane production to rely more on forced labour than family farms, thus impeding the development of a middle class and fostering inequality.

We also use the share of domestic credit to the private sector over GDP (dcredit) as our second instrument variable for inequality. The rationale behind this instrument reflects the theoretical argument put forward by Bénabou (2000) that in the context of capital market imperfections, access to credit and investment opportunities vary substantially among individuals with differential capital endowments, and that consequently lead to a persistence in income inequality.[9]

Finally, we follow the argument put forward by Aiyar and Ebeke (2020), and consider the adolescent fertility rate (adolfert) as our third instrumental variable. High fertility rates among adolescents are likely to adversely affect human capital endowments and future earnings, which in turn would worsen income inequality. Since higher adolescent fertility rates are likely to be more prevalent among low-income households, we use the lagged values of this indicator as an instrument to inequality.

3.2.4. Other controls

As for the economic and socio-demographic controls, we employed data from the World Bank's World Development Indicators (WDI) as our primary data source (World Bank, 2019). Data on institutional dimensions are drawn from the International Country Risk Guide (ICRG) dataset (ICRG 2018), which is published annually by the PRS Group.[10]

3.2.5. Study coverage

The present study covers 116 countries, 27 of which are in the SSA region, over the period 1990–2015.[11] All the variables used in the analysis are averaged over five-year periods.[12] This choice is motivated by the fact that comparable annual data for inequality are available only for a limited number of countries and by the evidence that inequality is a highly persistent variable. Furthermore, averaging data over time intervals makes the results less sensitive to the possibility of short-term fluctuations. Table A3 presents the summary statistics for all the variables used in the analysis.

On average, over the examined period, total government revenues represent nearly 23% of GDP at the global level. This share is lower for SSA (see Table A4), for which total revenues amount to approximately 17% of GDP. As for income inequality, the average value for the Gini index is about 45 points on a 0–100 scale. Compared to the global average,

[9] This instrumental variable has been used by previous studies (e.g., De Mello and Tiongson (2006) that empirically examine the causal relationship between inequality and redistribution.

[10] We are aware of the heterogeneity in the quality of data for the different groups of countries included in the analysis. We have relied on the most accurate, harmonized and comprehensive data sources available for cross-country analysis. Nevertheless, we acknowledge the possibility of having problems of measurement error due to data constraints.

[11] The list of countries included in the sample is referred to in Table A2.

[12] Variables' definitions and data sources are reported in Table A1.

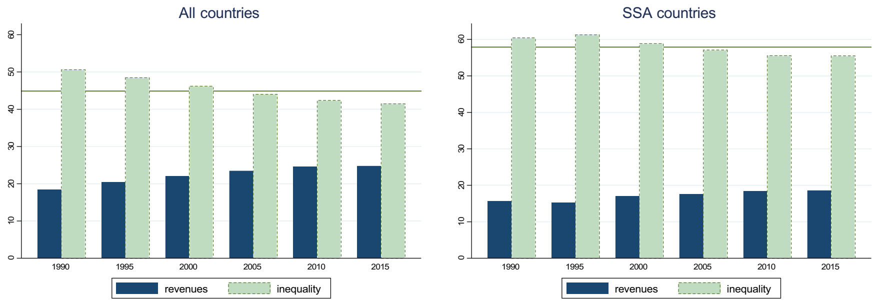

Figure 1. Total Revenues (GDP Share, %) and Inequality (Gini index)

SSA countries are characterized by a much higher level of inequality, with a mean value of almost 58 points. Figure 1 provides a general picture of the pattern characterizing the two main variables of interest over the reference period.[13] On a global level, the share of total government revenues over GDP shows an increasing pattern while income inequality exhibits a sizable reduction over the same period. In the case of SSA, we observe a similar pattern, although the trends in both total revenues and inequality are not strictly monotonic, especially with reference to the first decade.

In order to have a more detailed representation of the structure of total government revenues in the SSA region, we show in Figure B1 the average values of revenues and inequality by country. The next section presents the results of the econometric analysis.

4. Results

We begin the discussion by presenting the results of Model (1) based on a 'naive' pooled OLS estimator, which relies on the exogeneity assumption of inequality. The results in Table 1 (column 1) show a negative coefficient for the Gini index, indicating that higher levels of income inequality are associated with lower revenue capacity, thus acting as a detrimental factor in countries' resource mobilization efforts.

As discussed earlier, we suspect the OLS estimators to be biased, as the level of income inequality is unlikely to be independent from redistribution decisions, measured by total government revenues. In such a case, the unobservable error term would be correlated with the Gini index and the OLS would produce inconsistent parameter estimates. Therefore, we adopt an instrumental variable approach.

As shown in Table 1, we first compute Model (1) as an exactly identified model (columns 2 and 3), with the share of land used to produce wheat, relative to the share of land used for sugarcane production (wheatsugar) as the instrumental variable. We then compute the same Model (1) but with a richer set of instruments (columns 4 and 5), adding to wheatsugar two additional instruments: the share of domestic credit to the private sector over GDP (dcredit) and the adolescent fertility rate (adolfert). After conducting an endogeneity test, we find that the Gini index that measures the level of income inequality is in fact endogenous to redistribution in the specified model.[14] Therefore, we focus on the 2SLS estimators (columns 2–5), which provide consistent parameter estimates of the causal effect of inequality on redistribution.

Before turning our attention to the results, we test the validity of the IV procedure. First, we perform an under-identification test to assess the relevance of the instruments. A rejection

13 A slightly different view on the association between total revenues and inequality is provided by Figure A1, where country-period observations are plotted instead of the average values.

14 The null hypothesis assumes the regressor to be exogenous. Test results reject the null at a 5% level.

Table 1. Inequality and Total Government Revenues

	OLS	2SLS		2SLS	
Depvar	(1) Revenues	(2) Revenues	(3) Ineq	(4) Revenues	(5) Ineq
gini	−0.358***	−0.960***	–	−0.874***	–
	(0.122)	(0.293)		(0.259)	
yPPP	0.059	−0.032	−0.118***	−0.019	−0.126***
	(0.062)	(0.079)	(0.028)	(0.075)	(0.031)
agric	−0.013***	−0.019***	−0.007***	−0.018***	−0.007***
	(0.004)	(0.005)	(0.002)	(0.004)	(0.002)
unempl	0.020***	0.024***	0.011***	0.024***	0.011***
	(0.004)	(0.006)	(0.003)	(0.005)	(0.003)
trade	0.001**	0.001**	−0.000	0.001**	−0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
depratio	0.000	0.003	0.004***	0.003	0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
femlabpart	−0.000	−0.000	−0.002	−0.000	−0.003
	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
popdens	−0.000***	−0.001***	−0.000	−0.001***	−0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
govstab	0.031*	0.040**	0.001	0.039**	0.004
	(0.016)	(0.016)	(0.010)	(0.016)	(0.010)
intconfl	0.008	0.001	−0.004	0.002	−0.004
	(0.014)	(0.016)	(0.007)	(0.016)	(0.007)
corrup	0.059**	0.039	−0.021*	0.042	−0.025**
	(0.028)	(0.026)	(0.012)	(0.026)	(0.012)
ethnt	−0.018	−0.015	−0.008	−0.015	−0.008
	(0.022)	(0.023)	(0.010)	(0.022)	(0.010)
wheatsugar			−1.599***		−1.439***
			(0.188)		(0.182)
dcredit			–		0.001*
					(0.000)
adolfert			–		0.001*
					(0.001)
Constant	3.413***	6.444***	4.863***	6.017***	4.937***
	(0.953)	(1.722)	(0.314)	(1.550)	(0.314)
Observations	530	530	530	530	530
R-squared	0.679	0.629		0.642	
Endog test p-val			0.014		0.049
K-P rk LM st. p-val			0.000		0.000
K-P rk Wald F st.			72.52		28.24
Hansen J p-val					0.265

Notes: Global sample. OLS and 2SLS estimators. Depvar columns (1), (2) and (4): total revenues (% GDP, ln). Depvar columns (3) and (5): inequality (gini, ln). Panel-clustered (country level) standard errors in parentheses. Period dummies included ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

of the null indicates that the model is identified. Second, we perform a weak-identification test to assess whether the instruments are strongly correlated with the endogenous regressor. A value of the F statistics above the critical values denotes that the correlation is not weak. Third, we compute the Hansen test of over-identifying restrictions. In this case, a rejection of the null casts doubt on the validity of the instruments. Overall, the performed tests show that the IV approach is the appropriate one to estimate the effect of inequality on redistribution.

Looking at the first-stage regressions in Table A5 in the Appendix, we find that the selected instruments are statistically significant. Specifically, the sign of the wheatsugar variable is the expected one, capturing the negative association between the relative abundance of land for growing wheat and inequality. A higher share of domestic credit to the private sector, instead, seems to have a detrimental distributive effect, exacerbating inequality. This indicates that capital market development seems to occur at the cost of higher income inequality. Finally, higher fertility rates among young women are found to be correlated with higher inequality, as postulated by the literature.

Turning to the main structural equation, we find that inequality has a negative effect on revenues. Since we enter equation (1) with a log–log specification, the coefficient of the Gini index can be interpreted as elasticities; i.e., the percentage change in total government revenues as the outcome of one percentage change in the levels of income inequality, *ceteris paribus*. More specifically, we find that an increase in the Gini index by 1% leads to a decrease in total government revenues by approximately 0.87–0.96%, depending on the choice of the instruments set.

Regarding other control variables, the size of the economy, measured by GDP per capita, is positive but statistically insignificant, indicating a weak relationship between economic development and revenue collection. Other structural indicators show that the sectoral composition of output is relevant for revenue mobilisation. For example, the share of agriculture over GDP has a negative and significant association with total government revenues while trade openness shows a positive and statistically significant, although very small, association.

The coefficient for the unemployment rate shows a positive and significant sign, which at first sight may not be in line with conventional theoretical expectations. Further analysis below, show that the results are driven by the presence of several middle-income countries in our sample, which are characterized by high level of unemployment and high values of total revenues over GDP, which is indicative of the procyclicality of business cycles among many developing countries as reported by Alesina *et al.* (2008) and Talvi and Vegh (2005).

Most socio-demographic factors appear not significant in their association with total revenues, with the only exception of population density that shows a small, negative and significant association with total revenues. While the results may appear counter intuitive, they are influenced by the presence in our global sample of a large number of middle-income countries in Asia, with high population density and low shares of government revenues over GDP, as well as a group of countries with very low population density and high shares of government revenues.

Finally, regarding the controls for institutional factors such as government stability, the level of corruption within the political system, the level of political violence and the presence of ethnic tensions show the expected sign in their coefficients, however, only the parameter coefficient that measures the ability of governments to implement policies show a significant correlation with revenue collection.

Given the significant heterogeneity in the global sample, we estimate the reference model with more homogeneous groups of countries, following the World Bank's country classification by income levels. In addition, we estimate the model for SSA as a whole (the region of interest in this study), and then divide the sub-sample into two groups of middle-income or low-income countries. This allows us to reduce the threat of unobserved heterogeneity in the relationship between inequality and redistribution in the SSA region. Results from the 2SLS estimators are reported in Table 2.

As already pointed out, looking at the estimated coefficients from the global sample, we find a significant negative effect of inequality on total revenues. Taking the global sample of countries as a benchmark, the magnitude of the inequality elasticity of redistribution increases to 1.45% when the sample is restricted to high-income countries while slightly decreases to 0.81% when the analysis is restricted to middle-income countries. The direction of the relationship is also negative but statistically insignificant for the case of low-income countries, partly due to the smaller sample of countries falling into that income classification.

Table 2. Inequality Effects on Total Government Revenues

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gini (ln)	−0.874***	−1.446***	−0.808***	−1.316	2.522**	1.719*	−1.958*
	(0.259)	(0.348)	(0.239)	(1.563)	(1.233)	(0.887)	(1.126)
yPPP	−0.019	−0.027	−0.127	−0.138	0.387**	0.260**	−0.186**
	(0.075)	(0.203)	(0.081)	(0.164)	(0.151)	(0.112)	(0.093)
agric	−0.018***	0.016	−0.025***	−0.010**	0.001	−0.014	−0.016**
	(0.004)	(0.037)	(0.006)	(0.005)	(0.008)	(0.012)	(0.006)
unempl	0.024***	0.006	0.025***	0.006	−0.002	−0.001	0.015
	(0.005)	(0.008)	(0.005)	(0.009)	(0.012)	(0.007)	(0.011)
trade	0.001**	−0.000	0.002***	0.002**	0.003***	0.004***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
depratio	0.003	0.017*	0.001	0.008*	0.015**	0.004	−0.003
	(0.002)	(0.010)	(0.003)	(0.004)	(0.008)	(0.006)	(0.006)
femlabpart	−0.000	−0.021	0.004	−0.024*	−0.022*	−0.006	−0.045***
	(0.003)	(0.014)	(0.003)	(0.013)	(0.011)	(0.012)	(0.009)
popdens	−0.001***	−0.000	−0.001***	−0.000	−0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
govstab	0.039**	−0.014	0.042***	−0.061	0.037	0.031	−0.062*
	(0.016)	(0.027)	(0.016)	(0.037)	(0.035)	(0.026)	(0.032)
intconfl	0.002	−0.023	0.005	0.023	0.018	0.028	0.021
	(0.016)	(0.038)	(0.019)	(0.026)	(0.026)	(0.028)	(0.031)
corrup	0.042	0.101***	0.017	0.107**	0.013	−0.109***	0.057
	(0.026)	(0.032)	(0.038)	(0.052)	(0.065)	(0.034)	(0.059)
ethnt	−0.015	−0.031	−0.056*	0.151*	−0.033	−0.057*	0.148**
	(0.022)	(0.032)	(0.030)	(0.078)	(0.050)	(0.033)	(0.062)
Observations	530	174	285	71	141	73	68
R-squared	0.642	0.306	0.541	0.495	0.665	0.780	0.418
Hansen J p-val	0.265	0.427	0.108	0.265	0.122	0.531	0.668
K-P rk LM st. p-val	0.000	0.003	0.005	0.448	0.094	0.198	0.315
K-P rk Wald F st.	28.24	17.06	17.16	0.905	1.907	1.687	2.351

Notes: 2SLS estimators. Depvar: total revenues (% GDP, ln). IV estimates. 2SLS pooled estimator. Panel-clustered (country level) standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

Surprisingly, we find that the sign of the parameter estimate for the Gini index is positive and statistically significant for SSA as a whole, and also for middle-income countries, in the order of 2.52 and 1.72, respectively, although it turns negative, −1.96, when we restrict the sample to low-income countries (see Table 2, columns 5–7).[15]

One possible interpretation is that higher levels of inequality create the incentives for governments to redistribute. Under competitive electoral systems, political power is better distributed than income, so the median voter would have the power to persuade elites to redistribute (Meltzer and Richard, 1981). As Alesina and Perotti (1996):360 argue: 'in the fiscal channel explanation, the level of government expenditure and taxation is the result of

[15] We note that due to a finite sample problem, the estimated coefficients for middle-income and low-income countries in sub-Saharan Africa are likely to be affected by a weak identification bias. In order to limit this problem, we reduce the number of over identifying restrictions by using two out of the three instruments (see Harding *et al.* (2016); Andrews *et al.* (2019) for a discussion on the finite sample bias and weak instrument issues). Moreover, as discussed below in Section 5.2, we estimate Model (1) using a Limited Information Maximum Likelihood (LIML) estimator, which has better small sample performance than 2SLS with weak instruments.

Table 3. Natural Resource Rents and Taxes on Income Profits and Capital Gains as Percentage of GDP

	Natural resources rents			Taxes on income, profits and capital gains		
Regions	**1990–1995**	**2000–2015**	**Var %**	**1990–1995**	**2000–2015**	**Var %**
Global	4.84	6.41	32.48	6.13	7.36	20.11
High-income countries	6.28	6.83	8.79	11.92	10.82	−9.26
Middle-income countries	5.10	7.24	41.99	5.19	5.67	9.06
Low-income countries	1.69	1.70	0.06	2.09	2.80	33.95
SSA	4.87	6.09	25.05	4.49	4.89	8.77
SSA (MICs)	8.23	10.55	28.13	6.89	6.84	−0.76
SSA (LICs)	1.69	1.73	2.15	2.09	2.88	37.77

Source: Authors' calculations, based on UNU-WIDER (2019a).

a voting process in which income is a main determinant of a voter's preferences; in particular, poor voters will favour high taxation'.

However, we believe that this channel is implausible, at least in the context of SSA, due to two important reasons: first, despite recent progress toward more competitive electoral systems, the region continues to be dominated by autocracies and electoral autocracies, where the median voter is less influential in redistribution decisions than elites, which are via lobbying groups and practices of corruption closely linked to government power (Bénabou, 2000; Stiglitz, 2012; Carter, 2016; Kroeger, 2020). Second, taxes on income, profits and capital gains have remained largely stagnated, and under a 5% level in terms of GDP since the 1990s. Among African middle-income countries, this share is slightly higher, about 7% of GDP, but this has not only remained stagnated but in fact declined between the 1990s and 2000s (see Table 3).

We posit that the most plausible mechanism for the positive relationship between inequality and total government revenues in SSA, especially among middle-income countries, relates to the composition of government revenue sources, and in particular, to the large and growing contribution of natural resource rents to government's budgets. Indeed, natural resource rents represent the largest source of revenue for governments in middle-income Africa, accounting for roughly one-tenth of national income, after having experienced rapid growth between 1990s and 2000s (see Table 3).

The abundance of natural resource rents can affect redistributive preferences and tax policy choices among opportunistic incumbents, as tax redistribution and non-tax redistribution face different political and economic costs (Menaldo, 2016). Tax revenues are subject to stronger opposition from voters than non-tax revenues, especially when non-tax revenues are dominated by a windfall of natural resource rents. In this sense, the presence of natural resources allow incumbents to bypass the interdependent preferences problem, insofar levying higher taxes on the richest is not a key element in redistribution and resource mobilisation strategies (Currie and Gahvari, 2008). Furthermore, natural resource rents can boost autocratic and rent-seeking behaviour, which militates against the bargaining power of the median voter (Torvik, 2002; Collier, 2010; Bjorvatn and Naghavi, 2011), and since the extractive industries are capital intensive, they exacerbate income inequality via capital accumulation and wages to skilled workers that are higher than those of the median voter (Addison and Roe, 2018). This in turn impact positively on government revenues.

4.1. Top incomes adjusted inequality estimates

So far, we have discussed the results based on a Gini index, which may be biased due to the omission of top incomes in household surveys. Since the income share going to the richest individuals can have a strong influence on the shape of the Lorenz curve and the Gini index,

Table 4. Top-Incomes Adjusted Gini Indices

Variable	Truncation point	Obs	Mean	Standard deviation	Min	Max
Gini index	$t = 1$	530	44.901	12.361	14.123	81.071
	$t = 0.9975$	530	48.218	14.515	14.435	92.703
	$t = 0.9950$	530	50.421	15.817	14.681	95.585
	$t = 0.9925$	530	52.424	16.889	14.909	96.152
	$t = 0.9900$	530	54.323	17.850	15.123	96.555

Notes: Global. When t is set equal to one, truncation is not considered in the estimation. As the truncation point falls, the non-response rate in household surveys increases. Estimates based on grouped data from UNU-WIDER (2019b).

Table 5. Top-Incomes Adjusted Gini Indices

Variable	Truncation point	Obs	Mean	Standard deviation	Min	Max
Gini index	$t = 1$	141	57.914	8.080	45.690	81.071
	$t = 0.9975$	141	63.369	10.355	48.216	92.703
	$t = 0.9950$	141	66.964	11.253	49.792	95.585
	$t = 0.9925$	141	70.157	11.567	51.409	96.152
	$t = 0.9900$	141	73.115	11.672	53.090	96.555

Notes: SSA. When t is set equal to one, truncation is not considered in the estimation. As the truncation point falls, the non-response rate in household surveys increases. Estimates based on grouped data from the WIID (2019).

as well as on governments' redistributive decisions, we are interested in assessing the extent to which the impact of income inequality on government revenues changes by alternative assumptions on the shape of the income distribution. Therefore, we re-estimate the reference equation (1) with an alternative series of the Gini index, which is adjusted by the effect of top-incomes on the income distribution, based on specific assumptions about the truncation points that occur at the top percentiles as described earlier.

Before discussing the results, we present a summary statistics of the top-incomes adjusted Gini indices in Tables 4 and 5. As expected, we observe that the Gini index displays its lowest value when it is assumed that the distribution of income is not truncated, i.e., at $t = 1$. In contrast, when we assume that household survey data upon which the Gini indices are estimated are representative of the bottom 99% of the income distribution, i.e., with a truncation that excludes the richest 1%, a much higher level of income inequality is observed. Truncation points lying within such a range are associated with intermediate monotonic values of the Gini index.

The increase in the level of income inequality after adjusting for the effects of top incomes is particularly striking for the case of SSA, for which the mean value of the Gini index goes from 57.91 with no top-incomes adjustment, up to 73.12 when the income distribution is adjusted based on a truncation at the 0.99 percentile.

We present in Table 6 the results of the re-estimated equation (1), using the top-incomes adjusted Gini indices. We find that the size effect of income inequality on total government revenues is somehow contained, although marginally, when we account for the effect of top incomes.[16] The findings suggest that despite the very considerable impact that the richest individuals have on the shape of the income distribution, their inclusion in the estimates have a very small mitigating income inequality effect on total government revenues.

For the global sample, the negative inequality elasticity of government revenues goes down from −0.87 (with no truncation) to −0.81 (with a $t = 0.9900$), which seems to indicate

16 See comparatively baseline estimates in Table 2 and top-incomes adjusted estimates in Table 6.

Table 6. Inequality Effects on Total Government Revenues (Top-incomes Adjusted Gini Indices)

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gini index (ln), $t = 0.9975$	−0.851***	−1.399***	−0.795***	−0.573	2.351*	1.425*	−0.946
	(0.250)	(0.319)	(0.233)	(0.907)	(1.318)	(0.728)	(0.917)
Observations	530	174	285	71	141	73	68
R-squared	0.634	0.315	0.524	0.540	0.618	0.774	0.515
Hansen J *p*-val.	0.279	0.418	0.119	0.234	0.096	0.591	0.306
K-P rk LM st. *p*-val.	0.000	0.003	0.006	0.175	0.135	0.213	0.103
K-P rk Wald *F* st.	27.42	17.30	14.96	2.163	1.279	1.575	3.324
Gini index (ln), $t = 0.9950$	−0.834***	−1.372***	−0.787***	−0.526	2.369*	1.408*	−0.802
	(0.244)	(0.305)	(0.230)	(0.784)	(1.403)	(0.741)	(0.782)
Observations	530	174	285	71	141	73	68
R-squared	0.631	0.319	0.518	0.534	0.590	0.764	0.516
Hansen J *p*-val.	0.297	0.411	0.130	0.231	0.101	0.617	0.309
K-P rk LM st. *p*-val.	0.000	0.003	0.006	0.153	0.188	0.226	0.087
K-P rk Wald *F* st.	27.35	17.32	14.02	2.598	1.083	1.479	4.005
Gini index (ln), $t = 0.9925$	−0.819***	−1.348***	−0.778***	−0.518	2.388	1.489*	−0.724
	(0.239)	(0.295)	(0.228)	(0.718)	(1.468)	(0.813)	(0.706)
Observations	530	174	285	71	141	73	68
R-squared	0.631	0.322	0.517	0.529	0.576	0.755	0.516
Hansen J *p*-val.	0.313	0.403	0.140	0.233	0.099	0.677	0.314
K-P rk LM st. *p*-val.	0.000	0.003	0.006	0.146	0.212	0.222	0.082
K-P rk Wald *F* st.	27.49	17.31	13.71	2.922	1.068	1.488	4.426
Gini index (ln), $t = 0.9900$	−0.805***	−1.326***	−0.767***	−0.521	2.366	1.592*	−0.683
	(0.235)	(0.286)	(0.225)	(0.681)	(1.492)	(0.900)	(0.661)
Observations	530	174	285	71	141	73	68
R-squared	0.633	0.324	0.520	0.526	0.579	0.747	0.516
Hansen J *p*-val.	0.322	0.396	0.145	0.235	0.092	0.709	0.320
K-P rk LM st. *p*-val.	0.000	0.003	0.005	0.149	0.208	0.215	0.083
K-P rk Wald *F* st.	27.75	17.35	13.75	3.089	1.150	1.499	4.662

Notes: 2SLS estimators. Depvar: total revenues (% GDP, ln). IV estimates. 2SLS pooled estimator. Panel-clustered (country level) standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

that the contribution of top income earners to government revenues, via taxes on income and capital gains, may contain the negative relationship between the Gini index and government revenues, but just marginally.

In the case of SSA, this relationship becomes less elastic, from 2.52 to 2.37 in the region as a whole, and, from 1.72 and −1.96 to 1.59 and −0.68, for the cases of middle-income and low-income countries, respectively.[17] Thus, despite the very large effect of top incomes on income inequality in the SSA region, accounting for the richest does not lead to a sizable increase in government revenues. This may be explained by at least two important considerations. First, there is a limited scope for taxes on income, profits and capital gains to contribute to government revenues, partly because of the persistence of

[17] We note that the statistical significance of the parameter estimates for the full sample of sub-Saharan African countries and the sub-sample of low-income countries, disappears when accounting for the effects of top incomes.

informality and subsistence agriculture across the region.[18] Indeed, the share of income taxes to GDP had remained under a 5% level in SSA since the 1990s until recently, when it increased marginally. Among middle-income countries, that share is slightly higher, about 7%, although it has not changed since the 1990s, and in fact declined by about one percentage between the 1990s and the 2000s. In addition, the presence of tax incentives, and loopholes in double-taxation agreements facilitate tax evasion and tax avoidance in the extractive industries, which limits the revenue mobilisation capacity of natural resource rich countries (Ndikumana and Boyce, 2022; UNESCA, 2019).

The second consideration is in the domain of political economy. In the African context, characterised by imperfect competitive electoral systems dominated by elites, the effect of the median voter on redistribution is likely to be contained by the power of politically cohesive elites that have strong ties with incumbents and systems of patronage and clientelism (Acemoglu *et al.*, 2011). Thus, the preferences of the median voter are likely to be overshadowed by those privileged actors in society that shape policy processes and limit progressive fiscal reforms (Bardhan and Mookherjee, 2000). Consequently, the presence of high income inequality and even higher due to top incomes, would lead to a constrained redistribution, which is reinforced by the presence of natural resource rents as discussed earlier.

5. Robustness checks

In order to assess the reliability of our results, we perform a number of robustness checks. First, we estimate the reference model over comparable samples in terms of number of observations, by including dummies for the different country groups as well as their interactions with the inequality variable. Second, we use alternative estimators, specifically the two-step feasible generalised method of moments (GMM) and the limited- information maximum likelihood (LIML). Third, we apply in a random-effect panel estimator, which allows us to take into account the potential presence of unobserved individual effects.[19]

5.1. Model with interaction terms

The reference model relies on regional sub-samples, which limits the number of observations available for analysis, especially in the case of SSA. Therefore, in order to keep the sample of countries as large as possible, we extend model (1) by including a dummy variable that identifies country subgroups (CCd_i) considered in Table 2 and their interaction with income inequality ($I_{it} \times \text{CCd}_i$), and which takes the following form:[20]

$$R_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \beta_3 \text{CCd}_i + \beta_4 \left(I_{it} \times \text{CCd}_i \right) + \upsilon_t + c_{it}, \tag{4}$$

where β_1 denotes the marginal effect of income inequality for those countries which do not belong to the referred group, β_4 captures the difference in the relationship of interest (i.e., the effect of inequality on total government revenues) between the referenced group of countries and the rest of the world, while $\beta_1 + \beta_4$ measures the marginal effect of income inequality on total government revenues for the referenced group of countries. To illustrate, when

[18] Informal employment in the represents about 80–90% of total non-agriculture employment in low and lower-middle-income countries, whereas employment in agriculture, measured as percentage of total employment, remains above 60% in low-income countries and about 40% in lower-middle income countries (Niño-Zarazúa, 2019).

[19] We have also considered the possibility of applying a fixed-effect panel estimator, however, given the relevance of time invariant and persistent variables in our model, and that the use of a fixed-effect estimator would have limited the extension of the model to include country-group dummy variables and their interactions with inequality, we decided not to proceed further.

[20] These country subgroups are high-income, middle-income, low-income, countries in the sub-Saharan African region, and middle-income and low-income countries in that region.

Table 7. Inequality Effects on Total Government Revenues (Model with Interactions)

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gini (ln)	−0.874***	−0.762***	−1.013***	−0.910***	−1.098***	−1.116***	−0.921***
	(0.259)	(0.254)	(0.337)	(0.242)	(0.300)	(0.295)	(0.267)
yPPP	−0.019	0.049	0.012	−0.063	0.014	−0.027	−0.062
	(0.075)	(0.069)	(0.075)	(0.067)	(0.071)	(0.067)	(0.070)
agric	−0.018***	−0.016***	−0.015***	−0.017***	−0.012**	−0.015***	−0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
unempl	0.024***	0.022***	0.021***	0.023***	0.011**	0.012**	0.023***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
trade	0.001**	0.001*	0.001*	0.001**	0.001*	0.001*	0.001**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
depratio	0.003	0.002	0.003	0.003	0.003	−0.000	0.003
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
femlabpart	−0.000	0.001	0.002	0.003	−0.003	−0.003	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
popdens	−0.001***	−0.000***	−0.000***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
govstab	0.039**	0.032*	0.032**	0.028	0.042**	0.023	0.026
	(0.016)	(0.017)	(0.016)	(0.018)	(0.017)	(0.016)	(0.019)
intconfl	0.002	0.004	−0.000	−0.006	−0.007	−0.013	−0.005
	(0.016)	(0.015)	(0.015)	(0.016)	(0.019)	(0.017)	(0.017)
corrup	0.042	0.054*	0.059**	0.056**	0.035	0.044*	0.055**
	(0.026)	(0.028)	(0.028)	(0.027)	(0.026)	(0.025)	(0.027)
ethnt	−0.015	−0.013	−0.005	−0.002	−0.010	0.002	−0.004
	(0.022)	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)	(0.025)
CCd	–	0.230	−0.819	10.214	−12.528***	−7.325**	13.472
		(1.471)	(1.459)	(6.230)	(4.804)	(3.445)	(9.323)
CCd × gini	–	−0.108	0.261	−2.621*	3.118***	1.884**	−3.433
		(0.401)	(0.385)	(1.561)	(1.171)	(0.848)	(2.341)
Observations	530	530	530	530	530	530	530
R-squared	0.642	0.660	0.665	0.612	0.611	0.671	0.577
Hansen J p-val.	0.265	0.208	0.222	0.120	0.436	0.177	0.039
K-P rk LM st. p-val.	0.000	0.002	0.000	0.365	0.114	0.027	0.290
K-P rk Wald F st.	28.24	8.832	22.42	1.827	3.339	7.557	1.619
Linear combinat.: gini + (CCd × gini)		−0.870***	−0.751***	−3.530**	2.020**	0.768	−4.354*
		(0.309)	(0.252)	(1.530)	(1.016)	(0.673)	(2.320)

Notes: 2SLS estimators. Depvar: total revenues (% GDP, ln). IV estimates. 2SLS pooled estimator. Panel-clustered (country level) standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

looking at SSA, the coefficient β_1 will capture the effect of income inequality on government revenues for countries which do not belong to SSA, β_4 will measure the difference between sub-Saharan African countries and the rest of the world, whereas the linear combination $\beta_1 + \beta_4$ will measure effect of income inequality on total government revenues in SSA. Results of the model including the interactions are presented in Table 7. Overall, the findings from the model with interactions confirm previous results from the baseline model.

5.2. Alternative estimators

In order to mitigate the weak instrument problem in some specifications, we estimate the reference model by using alternative estimators. This step is motivated by the fact that the

Table 8. Inequality Effects on Total Government Revenues

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gini (ln)	−0.709***	−1.472***	−0.757***	−2.012	3.054***	1.641*	−2.251**
	(0.237)	(0.328)	(0.232)	(1.490)	(1.184)	(0.878)	(0.896)
yPPP	0.028	−0.013	−0.085	−0.194	0.406***	0.260**	−0.212***
	(0.069)	(0.181)	(0.078)	(0.151)	(0.151)	(0.112)	(0.071)
agric	−0.017***	0.020	−0.025***	−0.012***	0.003	−0.015	−0.018***
	(0.004)	(0.034)	(0.006)	(0.004)	(0.008)	(0.011)	(0.006)
unempl	0.024***	0.009	0.024***	0.006	−0.005	−0.002	0.016
	(0.005)	(0.007)	(0.005)	(0.009)	(0.012)	(0.007)	(0.011)
trade	0.001**	−0.000	0.002***	0.002***	0.003***	0.005***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
depratio	0.002	0.015**	0.001	0.008**	0.016**	0.003	−0.004
	(0.002)	(0.008)	(0.003)	(0.004)	(0.008)	(0.006)	(0.005)
femlabpart	−0.000	−0.023*	0.004	−0.019	−0.024**	−0.007	−0.047***
	(0.003)	(0.013)	(0.003)	(0.013)	(0.011)	(0.011)	(0.008)
popdens	−0.001***	−0.000	−0.001***	0.000	−0.001	0.001	0.002*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
govstab	0.039**	−0.010	0.042***	−0.058	0.048	0.025	−0.067**
	(0.016)	(0.027)	(0.015)	(0.036)	(0.035)	(0.024)	(0.028)
intconfl	0.010	−0.032	0.012	0.041*	0.016	0.031	0.023
	(0.015)	(0.034)	(0.018)	(0.022)	(0.026)	(0.027)	(0.030)
corrup	0.032	0.099***	0.029	0.082*	0.010	−0.108***	0.054
	(0.025)	(0.031)	(0.038)	(0.044)	(0.064)	(0.034)	(0.058)
ethnt	−0.017	−0.043	−0.060**	0.178**	−0.057	−0.055*	0.163***
	(0.022)	(0.031)	(0.030)	(0.074)	(0.047)	(0.033)	(0.052)
Observations	530	174	285	71	141	73	68
R-squared	0.660	0.277	0.539	0.338	0.594	0.784	0.350
Hansen J p-val	0.265	0.427	0.108	0.265	0.122	0.531	0.668
K-P rk LM st. p-val	0.000	0.003	0.005	0.448	0.094	0.198	0.315
K-P rk Wald F st.	28.24	17.06	17.16	0.905	1.907	1.687	2.351

Notes: GMM2S estimators. Depvar: total revenues (% GDP, ln). GMM2S pooled estimator. Panel-clustered (country level) standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

2SLS estimator can be biased in small samples and the bias can be worsen in the presence of over-identifying restrictions. We considered alternative estimators that are asymptotically equivalent to 2SLS but have better finite-sample properties.

We first adopt a two-step efficient generalized method of moments (GMM) estimator. Its higher efficiency compared to the 2SLS estimator derives from the use of an optimal weighting matrix, the over-identifying restrictions of the model, and the relaxation of the i.i.d. assumption. Results are presented in Table 8. In addition, we adopt a limited-information maximum likelihood estimator, which performs better than 2SLS in presence of weak instruments. Results are presented in Table 9. All in all, the findings from these alternative estimators confirm the results from the 2SLS model.

5.3. Alternative panel methods

As a third robustness check, we estimate the reference model based on a random-effect, instrumental variable (RE-IV) panel estimator, which takes into account the presence of unobserved individual effects in the error term. The reference model (1) can be specified as

Table 9. Inequality Effects on Total Government Revenues

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
gini (ln)	−0.901***	−1.526***	−0.868***	−3.869	4.119	1.782*	−2.026*
	(0.269)	(0.373)	(0.263)	(8.899)	(2.943)	(0.949)	(1.179)
yPPP	−0.023	−0.044	−0.128	−0.413	0.494*	0.263**	−0.192**
	(0.076)	(0.213)	(0.082)	(0.921)	(0.257)	(0.115)	(0.096)
agric	−0.018***	0.015	−0.025***	−0.015	0.009	−0.013	−0.016**
	(0.004)	(0.038)	(0.006)	(0.019)	(0.016)	(0.012)	(0.007)
unempl	0.024***	0.006	0.025***	0.013	−0.015	−0.001	0.016
	(0.005)	(0.008)	(0.005)	(0.029)	(0.025)	(0.007)	(0.012)
trade	0.001**	−0.000	0.002***	0.001	0.003*	0.004***	0.002**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
depratio	0.003	0.018*	0.001	0.004	0.020	0.004	−0.003
	(0.002)	(0.010)	(0.003)	(0.014)	(0.014)	(0.007)	(0.007)
femlabpart	−0.000	−0.021	0.004	−0.001	−0.028*	−0.006	−0.046***
	(0.003)	(0.015)	(0.003)	(0.081)	(0.017)	(0.012)	(0.009)
popdens	−0.001***	−0.000	−0.001***	0.000	−0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)
govstab	0.040**	−0.009	0.042***	−0.104	0.052	0.032	−0.062*
	(0.016)	(0.029)	(0.016)	(0.154)	(0.051)	(0.026)	(0.032)
intconfl	0.002	−0.024	0.004	0.025	0.011	0.027	0.021
	(0.016)	(0.039)	(0.019)	(0.051)	(0.038)	(0.029)	(0.031)
corrup	0.041	0.098***	0.021	0.055	0.042	−0.110***	0.054
	(0.026)	(0.032)	(0.039)	(0.209)	(0.099)	(0.035)	(0.060)
ethnt	−0.015	−0.030	−0.057*	0.238	−0.078	−0.057*	0.150**
	(0.023)	(0.033)	(0.030)	(0.322)	(0.094)	(0.033)	(0.063)
Observations	530	174	285	71	141	73	68
R-squared	0.638	0.258	0.530	−0.440	0.413	0.776	0.404
Hansen J *p*-val.	0.269	0.432	0.113	0.534	0.187	0.534	0.668
K-P rk LM st. *p*-val.	0.000	0.003	0.005	0.448	0.094	0.198	0.315
K-P rk Wald *F* st.	28.24	17.06	17.16	0.905	1.907	1.687	2.351

Notes: LIML estimators. Depvar: total revenues (% GDP, ln). LIML pooled estimator. Panel-clustered (country level) standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

follows:

$$R_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 X_{it} + \upsilon_t + \eta_i + u_{it} \quad (5)$$

where η_i denotes the individual unobserved effects and u_{it} is the idiosyncratic error. In a RE-IV model, it is assumed a strict exogeneity of the individual term η_i in addition to the orthogonality with respect to the independent variables. Before moving onto the estimation of the RE-IV model, we implement a Breusch-Pagan test to formally assess the potential presence of unobserved individual effects. The results reject the null according to which the variance of the unobserved effect is zero.[21] Therefore, we proceed to implement the RE-IV estimator. Results are presented in Table 10. In addition, we estimate the RE-IV model with interactions, to keep the sample of countries as wide as possible. The results are presented in Panel B of Table 10. Overall, the results for the global sample as well as for the group of sub-Saharan African countries confirm the previous findings.

21 H_0: var(η_i) = 0. Chibar2(01) = 540.43 (p-value = 0.000).

Table 10. Inequality Effects on Total Government Revenues

	Global sample				SSA		
	All countries	By income level			All countries	By income level	
		High	Middle	Low		Middle	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PANEL A							
gini (ln)	−0.834**	0.732	−0.739	−1.316	2.474*	1.719	−4.111
	(0.325)	(1.221)	(0.457)	(1.864)	(1.282)	(1.053)	(2.574)
yPPP	−0.052	0.275	−0.013	−0.138	0.380**	0.260*	−1.055**
	(0.075)	(0.354)	(0.098)	(0.196)	(0.160)	(0.133)	(0.469)
agric	−0.021***	0.016	−0.023**	−0.010*	0.001	−0.014	−0.059**
	(0.005)	(0.018)	(0.009)	(0.005)	(0.008)	(0.014)	(0.025)
unempl	0.013***	0.010	0.014***	0.006	−0.001	−0.001	0.003
	(0.003)	(0.007)	(0.005)	(0.011)	(0.012)	(0.008)	(0.049)
trade	0.002***	0.000	0.003***	0.002*	0.003***	0.004**	0.009**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)
depratio	0.001	0.003	−0.000	0.008	0.015*	0.004	−0.003
	(0.002)	(0.005)	(0.004)	(0.005)	(0.008)	(0.008)	(0.022)
femlabpart	−0.006*	−0.011*	0.003	−0.024	−0.022*	−0.006	−0.075
	(0.003)	(0.007)	(0.003)	(0.016)	(0.012)	(0.014)	(0.068)
popdens	−0.001***	0.000	−0.001***	−0.000	−0.001	0.001	0.002
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.005)
govstab	0.020*	0.002	0.028**	−0.061	0.036	0.031	0.121
	(0.011)	(0.017)	(0.013)	(0.045)	(0.038)	(0.031)	(0.082)
intconfl	0.018*	0.015	0.014	0.023	0.019	0.028	0.045
	(0.010)	(0.013)	(0.015)	(0.032)	(0.028)	(0.033)	(0.048)
corrup	0.025	0.035*	0.029	0.107*	0.014	−0.109***	−0.111
	(0.017)	(0.018)	(0.025)	(0.062)	(0.069)	(0.041)	(0.123)
ethnt	−0.018	−0.030	−0.022	0.151	−0.032	−0.057	−0.140
	(0.016)	(0.032)	(0.018)	(0.094)	(0.052)	(0.039)	(0.123)
Observations	530	174	285	71	141	73	68
Number of countries	116	41	61	14	27	14	13
Hansen J p-val.	0.504	0.347	0.407	.	0.137	.	.
K-P rk LM st. p-val.	0.000	0.641	0.010	0.448	0.097	0.198	0.415
K-P rk Wald F st.	10.233	0.529	5.960	0.905	1.900	1.687	0.839
PANEL B (model incl. interactions)							
gini (ln)		−0.733**	−1.002***	−0.761**	−1.024***	−1.119***	−0.802**
		(0.340)	(0.343)	(0.331)	(0.343)	(0.370)	(0.339)
CCd	–	0.195	−1.339	6.699**	−16.863**	−14.519**	7.945*
		(1.894)	(1.642)	(2.690)	(6.770)	(6.409)	(4.470)
CCd × gini	–	−0.091	0.388	−1.706**	4.173**	3.641**	−2.014*
		(0.523)	(0.434)	(0.677)	(1.657)	(1.587)	(1.128)
Observations		530	530	530	530	530	530
Number of countries		116	116	116	116	116	116
Hansen J p-val.		0.589	0.575	0.427	0.827	0.487	0.219
K-P rk LM st. p-val.		0.000	0.000	0.000	0.209	0.085	0.000
K-P rk Wald F st.		5.419	6.260	5.870	3.034	2.224	13.197
Linear combinat.:		−0.824*	−0.614*	−2.466***	3.148**	2.521*	−2.816**
gini + (CCd × gini)		(0.443)	(0.335)	(0.607)	(1.499)	(1.368)	(1.102)

Notes: RE-IV estimators. Depvar: total revenues (% GDP, ln). RE-IV panel estimator. Robust standard errors in brackets. Period dummies included. IVs col. (1)–(4): wheatsugar, adolfert, dcreditp. IVs col. (5)–(7): wheatsugar, dcreditp ***$p < 0.01$ **$p < 0.05$ *$p < 0.1$

6. Concluding remarks

The level of income inequality plays an important role in countries' economic performance and poverty reduction efforts. The literature has pointed out possible channels through which such relationships may operate. In the present study, we investigate the median voter hypothesis, by providing an empirical analysis of the relationship between income inequality and governments' revenue collection efforts.

In order to address the endogeneity of inequality, we have implemented a series of instrumental variable estimators and specifications, taking into account the panel structure of the available data, to test the validity of our results.

By looking at a wide sample of countries at the global level, we find a negative relationship between inequality and total government revenues, indicating that higher income inequality leads to a lower collection of government revenues. However, when we focus specifically on SSA, and subgroups of middle-income and low-income countries in the region, we observe a positive relationship, denoting that, higher income inequality leads to higher government revenues. Among the factors, which could be driving the result are the economic structure and sector composition of many African economies, especially for those middle-income countries which are rich in natural resources.

Similarly, another relevant issue is related to the composition of government revenues in most SSA countries, where the contribution of direct taxes is very limited.

Thus, the evidence suggests that it is not the median voter who through the power of persuasion in competitive electoral systems drives elites to redistribute via government revenues, but instead, it is the natural resource wealth of many African countries, which by allowing opportunistic incumbents to raise revenues without taxing the richest, exacerbate income inequality, which in turn impact positively on government revenues.

Acknowledgements

We are grateful to Théophile Azomahou, Gary Fields, Augustin K. Fosu, Njuguna Ndung'u, Yusi Ouyang, Abebe Shimeles, Erik Thorbecke, and seminar participants at the AERC meeting in November 2019 for helpful comments on earlier versions of this article. We are particularly grateful to the African Economic Research Consortium for financial support. Vanesa Jorda acknowledges partial financial support from the Fundación Ramón Areces (CISP20A6658).

Supplementary material

Supplementary material is available at *Journal of African Economies* online.

Data availability

The data and Stata do.files underlying this article are available in the article's online supplementary material.

References

Acemoglu D., Ticchi D., Vindigni A. (2011) 'Emergence and Persistence of Inefficient States', *Journal of the European Economic Association*, 9(2): 177–208.

Addison T., Roe A. (2018) *Extractive Industries: The Management of Resources as a Driver of Sustainable Development*. Oxford: Oxford University Press.

Adelman I., Robinson S. (1989) 'Income distribution and development', in H. Chenery and T.N. Srinivasan (eds), *Handbook of Development Economics*, vol. II, Chapter 19. Oxford: Elsevier.

Aghion P., Caroli E., Garcia-Penalosa C. (1999) 'Inequality and Economic Growth: The Perspective of the New Growth Theories', *Journal of Economic Literature*, 37(4): 1615–60.

Ahrens L. (2019) 'Unfair Inequality and the Demand for Redistribution: Why Not All Inequality Is Equal' *Socio-Economic Review*. 20(2): 463–87.

Aiyar S., Ebeke C. (2020) 'Inequality of Opportunity, Inequality of Income and Economic Growth', *World Development*, 136: 105115.

Akitoby B., Honda J., Miyamoto H., Primus K., Sy M. (2019) *Case Studies in Tax Revenue Mobilization in Low-income Countries*. IMF Working Paper 19/104. Washington, DC: International Monetary Fund (IMF).

Alesina A., La Ferrara E. L. (2005) 'Ethnic Diversity and Economic Performance', *Journal of Economic Literature*, 43(3): 762–800.

Alesina A., Perotti R. (1996) 'Income Distribution, Political Instability, and Investment', *European Economic Review*, 40(6): 1203–28.

Alesina A., Rodrik D. (1994) 'Distributive Politics and Economic Growth', *The Quarterly Journal of Economics*, 109(2): 465–90.

Alesina A., Baqir R., Easterly W. (1999) 'Public Goods and Ethnic Divisions', *The Quarterly Journal of Economics*, 114(4): 1243–84.

Alesina A., Tabellini G., Campante F. R. (2008) 'Why Is Fiscal Policy Often Procyclical?' *Journal of the European Economic Association*, 6: 1006–36.

Alvaredo F., Atkinson A. B., Piketty T., Saez E. (2013) 'The Top 1 Percent in International and Historical Perspective', *Journal of Economic Perspectives*, 27(3): 3–20.

Alvaredo F., Chancel L., Piketty T., Saez E., Zucman G. (2017) 'Global Inequality Dynamics: New Findings from WID.world'. *American Economic Review*, 107(5): 404–9.

Andrews I., Stock J., Sun L. (2019) 'Weak Instruments in IV Regression: Theory and Practice', *Annual Review of Economics*, 11: 727–53.

Atkinson A. B., Piketty T., Saez E. (2011) 'Top Incomes in the Long Run of History', *Journal of Economic Literature*, 49(1): 3–71.

Banerjee A. V., Newman A. F. (1993) 'Occupational Choice and the Process of Development', *Journal of Political Economy*, 101(2): 274–98.

Bardhan P. K., Mookherjee D. (2000) 'Capture and Governance at Local and National Levels', *American Economic Review*, 90(2): 135–9.

Bénabou R. (2000) 'Unequal Societies: Income Distribution and the Social Contract', *American Economic Review*, 90(1): 96–129.

Besley T., Persson T. (2014) 'Why Do Developing Countries Tax So Little?' *Journal of Economic Perspectives*, 28(4): 99–120.

Bird R. M., Martinez-Vazquez J., Torgler B. (2014) 'Societal Institutions and Tax Effort in Developing Countries', *Annals of Economics and Finance*, 15(1): 185–230.

Birdsall N., Haggard S. (2002) 'After the crisis: the social contract and the middle class in East Asia', in *When Markets Fail: Social Policy and Economic Reform*. New York: Russell Sage Foundation, pp. 58–96.

Bjorvatn K., Naghavi A. (2011) 'Rent Seeking and Regime Stability in Rentier States', *European Journal of Political Economy*, 27(4): 740–8.

Botlhole T., Asafu-Adjaye J., Carmignani F. (2012) 'Natural Resource Abundance, Institutions and Tax Revenue Mobilisation in Sub-Sahara Africa', *South African Journal of Economics*, 80(2): 135–56.

Boustan L., Ferreira F., Winkler H., Zolt E. M. (2013) 'The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970–2000', *Review of Economics and Statistics*, 95(4): 1291–302.

Breceda K., Rigolini J., Saavedra J. (2009) 'Latin America and the Social Contract: Patterns of Social Spending and Taxation', *Population and Development Review*, 35(4): 721–48.

Bussolo M., Davalos M.E., Peragine V., Sundaram R. (2019a) Toward a New Social Contract. *Taking on Distributional Tensions in Europe and Central Asia*. Washington, DC: The World Bank.

Bussolo M., Ferrer-i-Carbonell A., Giolbas A., Torre I. (2019b) *I Perceive therefore I Demand: The Formation of Inequality Perceptions and Demand for Redistribution. Policy Research Working Paper No*. 8926. Washington, DC: The World Bank.

Carter B. L. (2016) 'The Struggle over Term Limits in Africa: How International Pressure Can Help', *Journal of Democracy*, 27(3): 36–50.

Collier P. (2010) 'The Political Economy of Natural Resources', *Social Research*, 77(4): 1105–32.

Currie J., Gahvari F. (2008) ‘Transfers in Cash and in-Kind: Theory Meets the Data’, *Journal of Economic Literature*, 46(2): 333–83.
De Mello L., Tiongson E. R. (2003) *Income Inequality and Redistributive Government Spending*. Technical report, IMF Working Paper 314.Washington, DC: International Monetary Fund
De Mello L., Tiongson E. R. (2006) ‘Income Inequality and Redistributive Government Spending’, *Public Finance Review*, 34(3): 282–305.
Deininger K., Squire L. (1996) ‘A New Data Set Measuring Income Inequality’, *The World Bank Economic Review*, 10(1): 565–91.
Dioda L. (2012) *Structural Determinants of Tax Revenue in Latin America and the Caribbean, 1990–2009*. Technical report. Santiago de Chile, Chile: Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).
Drummond P., Daal W., Srivastava N., Oliveira L. E. (2012) *Mobilizing Revenue in Sub-Saharan Africa; Empirical Norms and Key Determinants*. Technical report. Washington, DC: International Monetary Fund.
Easterly W. (2007) ‘Inequality Does Cause Underdevelopment: Insights from a New Instrument’, *Journal of Development Economics*, 84(2): 755–76.
FAO (2019) ‘FAOSTAT *Crops* and *Livestock* Products’. Available at: FAO: FAOSTAT.org. Accessed on 15 June 2019.
Galor O., Zeira J. (1993) ‘Income Distribution and Macroeconomics’, *The Review of Economic Studies*, 60(1): 35–52.
Gisselquist R. M., Leiderer S., Niño-Zarazúa M. (2016) ‘Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational “Diversity Dividend”’, *World Development*, 78: 308–23.
Gnangnon S. K., Brun J.-F. (2019) ‘Trade Openness, Tax Reform and Tax Revenue in Developing Countries’, *The World Economy*, 42(12): 3515–36.
Habyarimana J., Humphreys M., Posner D. N., Weinstein J. M. (2007) ‘Why Does Ethnic Diversity Undermine Public Goods Provision?’ *American Political Science Review*, 101(4): 709–25.
Harding M., Hausman J., Palmer C. (2016) ‘Finite Sample Bias Corrected IV Estimation for Weak and Many Instruments’, *Advances in Econometrics*, 36: 245–73.
International Country Risk Guide, ICRG (2018). Table 3B: Researcher’s Dataset. *The PRS Group*.
Jenkins S. P. (2015) ‘World Income Inequality Databases: An Assessment of Wiid and Swiid’, *The Journal of Economic Inequality*, 13(4): 629–71.
Jorda V., Niño-Zarazúa M. (2019) ‘Global Inequality: How Large Is the Effect of Top Incomes?’ *World Development*, 123: 104593.
Khattry B., Rao J. M. (2002) ‘Fiscal Faux Pas?: An Analysis of the Revenue Implications of Trade Liberalization’, *World Development*, 30(8): 1431–44.
Kimenyi M. S. (2006) ‘Ethnicity, Governance and the Provision of Public Goods’, *Journal of African Economies*, 15(Supplement 1): 62–99.
Kroeger A. M. (2020) ‘Dominant Party Rule, Elections, and Cabinet Instability in African Autocracies’, *British Journal of Political Science*, 50(1): 79–101.
Kuznets S. (1955) ‘Economic Growth and Income Inequality’, *The American Economic Review*, 45(1): 1–28.
Leigh A. (2007) ‘How Closely Do Top Income Shares Track Other Measures of Inequality?’ *The Economic Journal*, 117(524): F619–33.
Lewis W. A. (1954) ‘Economic Development with Unlimited Supplies of Labour’, *The Manchester School*, 22: 139–91.
Li H., Zou H.-F. (1998) ‘Income Inequality Is Not Harmful for Growth: Theory and Evidence’, *Review of Development Economics*, 2(3): 318–34.
Lindert P. H. (1996) ‘What Limits Social Spending?’ *Explorations in Economic History*, 33(1): 1–34.
Lindqvist E., Östling R. (2013) ‘Identity and Redistribution’, *Public Choice*, 155(3–4): 469–91.
Luebker M. (2014) ‘Income Inequality, Redistribution, and Poverty: Contrasting Rational Choice and Behavioral Perspectives’, *Review of Income and Wealth*, 60(1): 133–54.
Lupu N., Pontusson J. (2011) ‘The Structure of Inequality and the Politics of Redistribution’, *American Political Science Review*, 105(2): 316–36.
McCarty N., Pontusson J. (2011) *The Political Economy of Inequality and Redistribution*. Oxford: Oxford University Press.
Meltzer A. H., Richard S. F. (1981) ‘A Rational Theory of the Size of Government’, *Journal of Political Economy*, 89(5): 914–27.

Menaldo V. (2016). *The Institutions Curse: Natural Resources, Politics, and Development*. Cambridge, UK: Cambridge University Press.
Milanovic B. (2000) 'The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data', *European Journal of Political Economy*, 16(3): 367–410.
Moene K. O., Wallerstein M. (2001) 'Inequality, Social Insurance, and Redistribution', *American Political Science Review*, 95(4): 859–74.
Ndikumana L., Boyce J.K. (eds.) (2022) *On the Trail of Capital Flight from Africa: The Takers and the Enablers*. Oxford: Oxford University Press.
Niño-Zarazúa M. (2019) 'Welfare and Redistributive Effects of Social Assistance in the Global South', *Population and Development Review*, 45(S1): 3–22.
Niño-Zarazúa M., Roope L., Tarp F. (2017) 'Global Inequality: Relatively Lower, Absolutely Higher', *Review of Income and Wealth*, 63(4): 661–84.
Ostry M. J. D., Berg M. A., Tsangarides M. C. G. (2014) *Redistribution, Inequality, and Growth*. Staff Discussion Note SDN/14/02. Washington, DC: International Monetary Fund.
Persson T., Tabellini G. (1994) 'Is Inequality Harmful for Growth?' *The American Economic Review*, 84(3): 600–21.
Piketty T., Saez E. (2013) 'Top Incomes and the Great Recession: Recent Evolutions and Policy Implications', *IMF Economic Review*, 61(3): 456–78.
Scervini F. (2012) 'Empirics of the Median Voter: Democracy, Redistribution and the Role of the Middle Class', *The Journal of Economic Inequality*, 10(4): 529–50.
Schwabish J.A., Smeeding T.M., Osberg L. (2006) 'Income distribution and social expenditures', in Papadimitriou, D.B. (ed.), *The Distributional Effects of Government Spending and Taxation*. London: Palgrave Macmillan, pp. 247–88.
Sen Gupta A. (2007) *Determinants of Tax Revenue Efforts in Developing Countries*. IMF Working Papers, No. 39. Washington, DC: International Monetary Fund.
Shelton C. A. (2007) 'The Size and Composition of Government Expenditure', *Journal of Public Economics*, 91(11): 2230–60.
Sokoloff K. L., Engerman S. L. (2000) 'Institutions, Factor Endowments, and Paths of Development in the New World', *Journal of Economic Perspectives*, 14(3): 217–32.
Stiglitz J. E. (2012) *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York: WW Norton & Company.
Talvi E., Vegh C. A. (2005) 'Tax Base Variability and Procyclical Fiscal Policy in Developing Countries', *Journal of Development Economics*, 78(1): 156–90.
Torvik R. (2002) 'Natural Resources, Rent Seeking and Welfare', *Journal of Development Economics*, 67(2): 455–70.
UNESCA (2019) *Economic Report on Africa 2019: Fiscal Policy for Financing Sustainable Development in Africa*. Addis Ababa, Ethiopia: United Nations Economic Commission for Africa.
UNU-WIDER (2019a) *Government Revenue Dataset*. Available at: UNU-WIDER: Government Revenue Dataset (GRD). Accessed on 15 January 2019.
UNU-WIDER (2019b) *World Income Inequality Database*. Available at: UNU-WIDER: World Income Inequality *Database* (WIID). Accessed on 20 January 2019.
World Bank (2019) *World Development Indicators*. Available at: World Development Indicators | DataBank (worldbank.org). Accessed on 20 January 2019.
Zoellick R.B. (2011) *The Middle East and North Africa: A New Social Contract for Development*. Speech delivered at the Peterson Institute for International Economics, Washington, DC, April 6, 2011. Washington, DC. World Bank.