



WIDER Working Paper 2016/133

Redistribution around the world

Causes and consequences

Markus Jäntti,¹ Jukka Pirttilä,² and Risto Rönkkö³

November 2016

Abstract: This paper re-examines the determinants and consequences of redistribution in light of improved data and methods relative to earlier literature. In particular, we use the latest version of the UNU-WIDER's Income Inequality Database to have the best available estimates of both pre- and post-redistribution inequality for the largest set of countries and periods. We also tackle head-on problems of related to model specification that risks generating large biases in estimates because of mechanical associations between variables. The paper also shows how earlier results on the determinants of redistribution and the impacts of redistribution on growth can be misleading.

Keywords: cross-country analysis, growth, inequality, redistribution

JEL classification: H11, O11, O47

Figures and tables: authors' own work.

Acknowledgements: We are grateful to seminar audiences at HECER–WIDER development lunch and the Summer Meeting of the Finnish Economists for useful comments. Support from the Academy of Finland (Grant 268082) is gratefully acknowledged.

¹University of Helsinki and VATT, Helsinki, Finland; ²UNU-WIDER, Helsinki, and University of Tampere, Finland; ³UNU-WIDER, Helsinki, and University of Tampere, Finland, corresponding author: risto@wider.unu.edu.

This study has been prepared within the UNU-WIDER project on 'World inequality'.

Copyright © UNU-WIDER 2016

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-177-2

Typescript prepared by the Authors and Anna-Mari Vesterinen.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

Inequality has in recent years become a key economic and social policy concern. Researchers have documented rising top income and wealth shares, people have protested in the streets about rising differences in income levels, and researchers and policymakers alike have started increasingly to worry about the potential negative consequences of inequality on social distress and economic performance.

Much of the discussion is quite vague about what type of inequality is being debated. Top income shares, for example, refer to market income (or factor income) inequality, whereas what should matter the most for welfare analysis is net income inequality (inequality of disposable income or even consumption inequality). This brings us to the centre stage the role of redistribution: the differences between market and net income inequality that stem from the effects of government fiscal policies, taxes and benefits. Much less is known about *redistribution* in the world than *inequality* in the world. This is regrettable since redistributive policies are actually the only direct and quick policy instrument that governments can use to try to curb disposable income inequality.¹

One key study, which also received a lot of media attention,² dealing with redistribution is Ostry, Berg, and Tsangarides (2014). In that paper the authors use the cross-country panel SWIID data to test whether inequality and redistributive actions by governments actually hamper growth (for information about the SWIID, see Solt 2013). This would be the case if the well-known efficiency-equity trade off were true at the macro level. They find that disposable income inequality is negatively associated with future growth, whereas redistribution (the difference between gross and net inequality) is not. This suggests that the efficiency-equity trade-off would not seem to hold.

One problem with the Ostry, Berg, and Tsangarides (2014) study, which the authors duly acknowledge themselves when cautioning against drawing strong policy recommendations from their work, is quality of the income inequality data they use. In developing countries, in particular, household surveys that are used to measure living standards are only undertaken approximately once every five to ten years, and for many years, there is no data on inequality for these countries. Also, often only one type of inequality measure is available. In the poorest countries, inequality typically refers to consumption inequality. Often, nothing is known about redistribution, which is the key variable that the Ostry, Berg, and Tsangarides (2014) analysis requires. The SWIID database (see Solt 2009, 2013) is an attempt to remedy the data-sparseness problem. The data are based on extensive imputations, where observations from the same countries in other years and other countries in the same year are used to impute both gross and net inequality indices for countries without the required data in a particular year. However, others have been very critical of the extensive use of imputations in general and the particular type of imputations the SWIID uses, and caution against using the dataset in econometric work (see Jenkins 2015a, and Ferreira, Lustig, and Teles 2015). They recommend using the WIID dataset, compiled and maintained by UNU-WIDER, instead.

This paper has three main goals. First, we explore how much actual, rather than simulated, data there are about redistribution at the country level. For this purpose, we utilize the WIID dataset. We also describe the extent and trends in redistribution. A key finding related to the first goal of the paper is that there are surprisingly and unfortunately few observations on redistribution. Second, using as much data that is reliable as possible, the paper examines what drives redistribution by examining the determinants of it using cross-country regression techniques. And third, the paper looks at a key potential consequence of redistribution, on growth.

¹This is not to say that gross income inequality could not be affected. It clearly can be, e.g. via educational policies that affect relative wage rates, but often such policies take a longer time span to take effect.

²See e.g. Economist (2014) or Guardian (2014)

Two main features characterize our approach. We take the data issues in cross-country analysis of inequality and redistribution seriously, as was strongly recommended by Atkinson and Brandolini (2001), and work with real, comparable observations, rather than imputations. Likewise, we tackle some of the econometric challenges that the earlier work on the determinants of redistribution has not addressed. In particular, we discuss the possible problems that arise from having the same variable – market income inequality in the redistribution measure which is the difference between market and net income inequality – as a left-hand side variable and as a right-hand side variable – where it is sometimes called inherent inequality. That is, we explore the consequences for estimation of the mechanical correlation of having the same variable on both sides of the regression equation.

Finally, a key part of the paper is to replicate the analysis of Ostry, Berg, and Tsangarides (2014) using the WIID instead of the SWIID data. It turns out that using actual rather than imputed data leads to different results regarding the impact of inequality on growth. But perhaps the main lesson is that when doing so, the sample size is radically reduced, and therefore it is very likely that we do not (at least yet) have the data available that one would need to carry out the analysis they set out to accomplish.

The paper proceeds as follows. Section 2 discusses the earlier literature on the determinants of inequality and some of the relevant work from the wide body of research on the impacts of inequality on growth. It also briefly summarizes the approach and key results from the Ostry, Berg, and Tsangarides (2014) study. Section 3 presents the data we use. We especially discuss the merits of using real rather than imputed data on inequality and redistribution. Section 4 discusses our econometric approach and the drawbacks in some of the earlier research that deals with factors explaining redistribution. Section 5 presents the results on the determinants of redistribution. It also replicates the results from the replication of the Ostry, Berg, and Tsangarides (2014) analysis. Section 6 concludes.

2 Literature review

Two sets of research work lay the analytical framework for thinking about income inequality and redistribution. One is provided by the Mirrlees (1971) optimal taxation theory and another by the median voter theory of Meltzer and Richard (1981). The analysis in Mirrlees bases on the redistributive preferences of society and highlights the classic trade-off in taxation between equality and efficiency, whereas Meltzer and Richard focus on the political process determining the level of redistribution.

Much of the empirical literature on the determinants of redistribution is focused on testing the median income voter theory. Meltzer and Richard develop a labour income tax model, in which individuals differ by their productivity and are provided lump-sum transfers financed by a flat rate tax. With a majority voting rule (and a hundred percent voter turnout) the tax rate, and consequently the extent of redistribution, is chosen by the median income voter. The authors show that as income inequality (measured by the difference between mean and median income) increases, the extent of redistribution voted for by the median-income voter increases.

These theories are usually tested³ by regressing a measure of redistribution on a gross-income inequality measure and control variables. Redistribution is in most cases operationalized as the difference between the market income and net income Ginis, but changes in pre- and post-redistribution income shares of income decile groups as well as social expenditure are also used. In most cases, authors estimate fixed-effects models using country-level panel data. The literature has brought two main conclusions. First, Tuomala and Tanninen (2005), Scervini (2012) among others find that pre-redistribution

³The most prominent studies include Karabarbounis (2011), Luebker (2014), Milanovic (2000, 2010), Tuomala and Tanninen (2005), and Scervini (2012).

income inequality leads to increased redistribution, suggesting that there is more demand or scope for redistribution in more unequal societies. The result emerges especially strongly from studies that use direct measures of redistribution (such as the difference between the pre- and post-redistribution Gini) as opposed to more indirect measures (such as social expenditure that measures the extent of total government expenditure, not only redistributive efforts) as a measure of redistribution. Second, the median income voter theory is rejected, as the median-income voter does not seem to hold a special role in decision making.

Another focal point of the inequality discussion is its effects on economic growth. It is often argued that equality-enhancing policies distort incentives, and are harmful for growth (Okun 1975). Alesina and Rodrik (1994), Bertola (1993), Perotti (1993), Persson and Tabellini (1994) argue that a high market income inequality causes redistributive policies, which are detrimental for growth. Income inequality can also promote growth if investment is curtailed by low level of income or wealth and there are growth enhancing externalities with investment. In that case, higher inequality allow the rich to accumulate a minimum amount of capital to invest growth enhancing activities (Barro 2000; Perotti 1993). On the other hand, inequality may deprive the poor of their ability to enhance their health or educate themselves, especially in the presence of credit constraints (Garcia-Penalosa, Caroli, and Aghion 1999). Income inequality may also cause socio-political conditions that are not conducive of growth (Alesina and Perotti 1996; Rodrik 1999).

Much of the empirical literature finds that inequality, in the medium term, hampers growth (Alesina and Rodrik 1994; Ostry, Berg, and Tsangarides 2014; Perotti 1996; Persson and Tabellini 1994). In addition to inequality, (Ostry, Berg, and Tsangarides 2014) consider also effects of redistribution on growth. Studying medium-term growth, they use SWIID Solt (2013) data with Generalized Method of Moments (GMM) estimation framework, and find that inequality hampers economic growth while there are no effects from redistribution on growth. SWIID data, however, contain a large number of imputed data points, which is a cause of concern when interpreting the results. To see if their results prevail, we replicate their estimations with WIID data, where all observations come from actual income inequality surveys.

3 Data

The source of our inequality data is the World Income Inequality Database, the WIID, version 3.3 (UNU-WIDER 2015),⁴ is secondary database for income inequality data. Aim of the WIID is to gather data for as many countries and years as possible, while documenting background information for data thoroughly. Depending on data source income inequality data differs from income and population concepts, sample sizes and statistical methods used. The background information helps database users to compare observations from different original sources for their statistical concepts and overall validity for research question in hand. Atkinson and Brandolini (2001), in their review and assessment of cross-country inequality analysis, pointed out that it is necessary to know to what type of inequality measures the indices refer to, and the WIID has been developed with this requirement in mind. The WIID also has a quality rating of the reported Gini indices.

The WIID data was assessed recently by Jenkins (2015b). He came to the conclusion that the WIID is a reliable source for cross-country work on inequality. He, and the authors of a synthesis chapter for the datasets reviewed (Ferreira, Lustig, and Teles 2015), both recommend using the WIID rather than

⁴The data set is open access and available at <https://www.wider.unu.edu/project/wiid-world-income-inequality-database>.

the SWIID.⁵ The main reason for preferring the WIID over the SWIID is that the SWIID is entirely based on imputations, whereas with the WIID, users utilize only actual, not simulated, data. Moreover, Jenkins (2015b) is critical of the particular type of imputations used in the SWIID, which are in his opinion remarkably complicated and opaque. Assuming constancy of ratios of Gini coefficients across data series within groups is not an innocuous assumption, and the smoothing of the series used may be excessive. Jenkins concludes that the SWIID implementation of imputation is not sufficiently credible.

Jenkins (2015b) also insists that users of the WIID make clear the algorithm they use for selecting the data. This is required as the data set has multiple observations for each country and year. We follow this recommendation and explain in detail the type of data selection mechanism used for our analysis in Appendix 6. The gist of the design is to calculate redistribution either as a difference between gross and net income inequality, or if this measure is not available, as a difference between gross income and consumption inequality. When doing so, we prefer high quality observations to lower quality observations and estimates covering the entire population and the whole country.

As inequality data for developing countries are only available for selected years, we take five-year averages (using all possible observations within that window), first five-year period being 1976-1980. Often the latest period is 2006-2010.

The rest of the data we use come from conventional sources. Data on GDP per capita, population and openness (share of exports plus imports of GDP) are from the World Bank World Development Indicators (2016) (WDI). In addition to the usual macroeconomic variables, we control for political and institutional factors. Regarding to electoral and governmental institutions, we include dummies for federal government system and plural electoral system. The latter refers to electoral systems in which voters cast a vote for a single candidate within a voting district, and the candidate who gets most votes wins the district. Data sources for the federal dummy is the International Monetary Fund and for the plurality dummy the Database of Political Institutions (Beck et al. 2001). We also consider the level of democracy or autocracy using the Polity2 variable from PolityIV project. The variable takes incremental values between -10 and 10 , with -10 referring the most autocratic form and government and 10 the most democratic government. We control for ethnic fractionalization with the Alesina et al. (2003) index which takes higher values for higher fractionalization. Lastly, we include a commodity export dummy for countries which have had average net exports of fuels, ores, or agricultural products more than 10 percent of GDP over the 1990s and 2000s. The source for the trade data is the WDI.

We present next descriptive material regarding the data on redistribution and recent trends in redistributive policies. Figure 1 shows how many observations on redistribution there are using five-year averaged data for the largest set of observations. It is already clear from this graph that the great majority of observations originate from high-income countries and hence examining redistribution in developing countries alone is challenging because of the paucity of data.

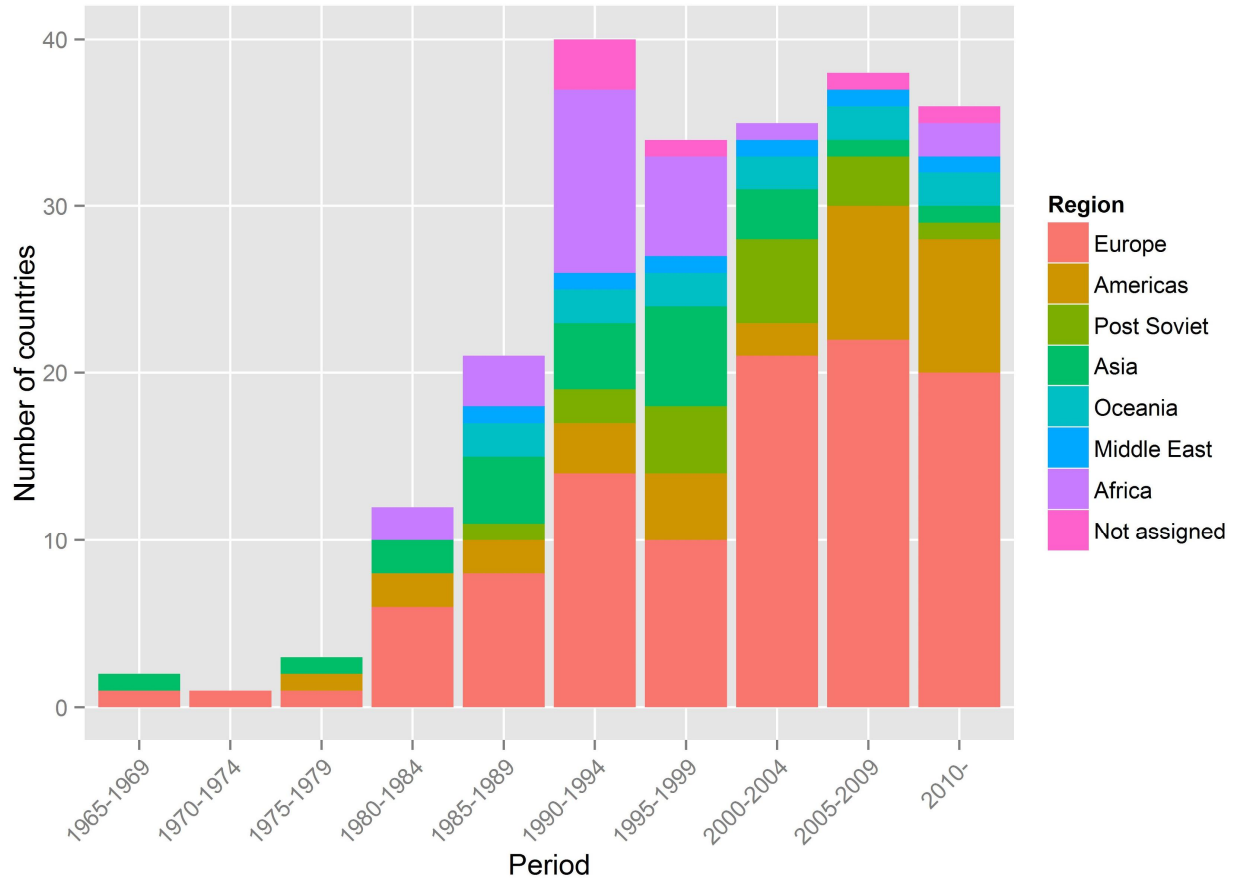
Figure 2 shows how the extent of redistribution is the highest in European countries, whereas the level is fairly low in many developing regions, including Asia. The limited availability of data needs to be kept in mind when interpreting the figure, however.

4 Models and methods

Studies that examine what drives redistribution typically regress either the *reduction* or the *relative reduction* in the Gini coefficient on moving from pre- to post-tax, post-transfer income. That is, re-

⁵Recall that Ostry, Berg, and Tsangarides (2014) use the SWIID.

Figure 1 Redistribution observations over countries.



searchers estimate a version of either (or both) of the regression equations

$$G_{it}^{pre} - G_{it}^{post} = \alpha + \beta G_{it}^{pre} + z'_{it} \delta + u_{it} \quad (1)$$

or

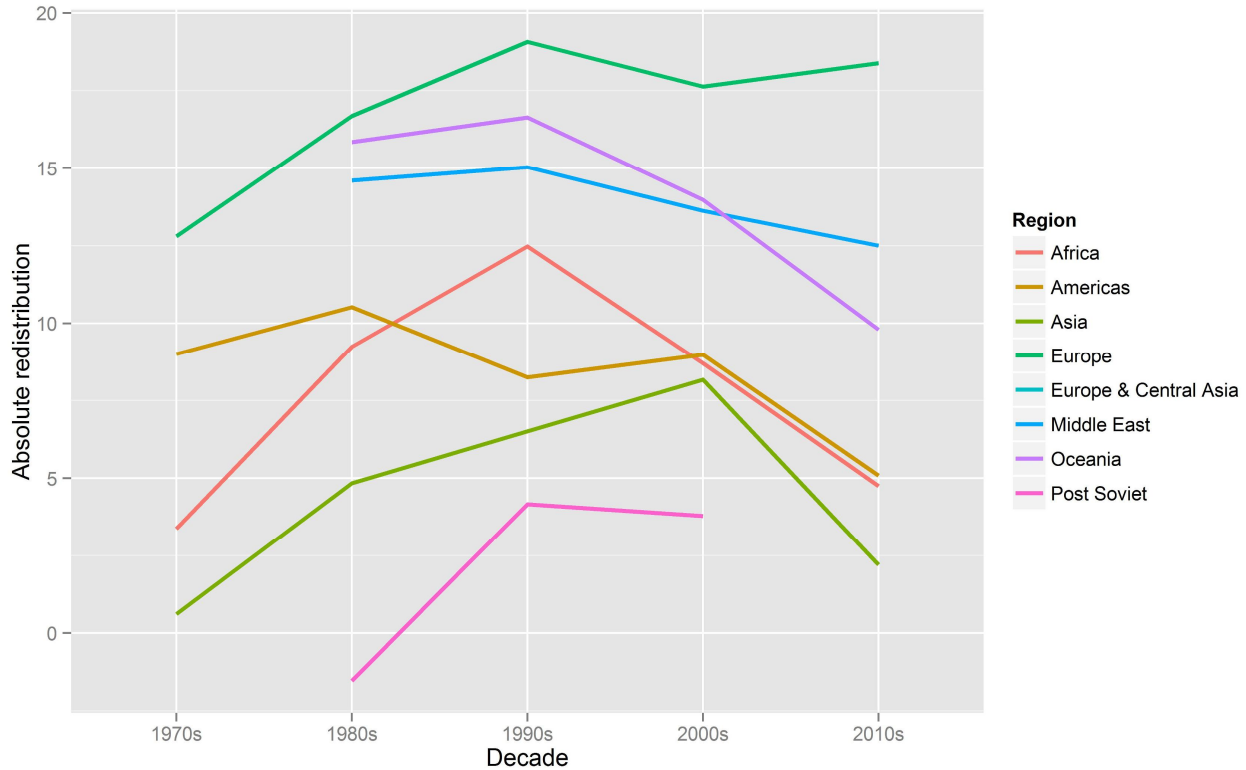
$$\frac{G_{it}^{pre} - G_{it}^{post}}{G_{it}^{pre}} = \alpha^* + \beta^* G_{it}^{pre} + z'_{it} \delta^* + v_{it} \quad (2)$$

where G is the Gini coefficient with pre and $post$ denoting market and disposable income, z is a vector of controls, u, v are error terms, and the unstarred and starred α, β, δ are parameters to be estimated. The indices indicate country i and time period t . The main objects of interest are β, β^* , which capture the extent to which redistribution varies with market income inequality, although several elements of δ, δ^* may also be of interest.

The inclusion of G^{pre} on both the left and right hands sides of the regression equations lead to several problems, however. Let us start with equation 1, which accounts for the (absolute) reduction in inequality. Ignoring, for now, the controls z , the intercept, and suppressing the indices, the linear projection of $G^{pre} - G^{post}$ on G^{pre} , i.e.,

$$\beta = \frac{\text{Cov}(G^{pre} - G^{post}, G^{pre})}{\text{Var}(G^{pre})} = \frac{\text{Cov}(G^{pre}, G^{pre})}{\text{Var}(G^{pre})} - \frac{\text{Cov}(G^{post}, G^{pre})}{\text{Var}(G^{pre})} = 1 - b, \quad (3)$$

Figure 2 The extent of absolute redistribution in different areas.



where b is the linear projection of G^{post} on G^{pre} .

Why is estimation of β from equation 1 a problem? Arguably, for at least two reasons. First, suppose there is *no* true association of market inequality with disposable income inequality; i.e., that b is zero. Then the estimate of β is unity, which tells us nothing about the association of market income inequality and redistribution. However, even if pre and post inequality are related (i.e., $b \neq 0$), inclusion of pre inequality on both the left and right hand sides leads to a serious endogeneity problem; i.e., by construction, u and G^{pre} are correlated. The simple solution to this problem is to estimate the linear projection of post on pre inequality, b , in a regression equation that also includes the controls z and work out the β this implies.

While all normal concerns about drawing causal inference apply, there is an additional, potential problem which needs to be treated with caution. Namely, the Gini coefficient is the expected mean difference divided by mean income (times two),

$$G = \frac{E(|Y_1 - Y_2|)}{2\mu}. \quad (4)$$

To simplify exposition, suppose you regress the natural logarithm of the Gini on the \ln of mean income,

$$\ln G = a + b\mu + e = \ln E(|Y_1 - Y_2|) - \ln \mu = a + b \ln \mu + e. \quad (5)$$

This ends up reproducing the mechanical association endogeneity problem described above. While mean disposable income is usually not one of the controls included in z , GDP per capita, which is closely linked to mean *market* income, usually is. We believe that this problem also is mitigated by the estimation of β from the the linear projection of post on pre inequality and control variables, because the association of disposable income with GDP per capita is less than that of market income.

What about *relative* redistribution, i.e., estimation of equation 2? Again, ignoring controls, the intercept and indices, the coefficient of interest is

$$\beta^* = \frac{\text{Cov}\left(\frac{G_{it}^{pre} - G_{it}^{post}}{G_{it}^{pre}}, G_{it}^{pre}\right)}{\text{Var}(G_{it}^{pre})}. \quad (6)$$

The numerator of this can be re-written as

$$\text{Cov}\left(\frac{G_{it}^{pre}}{G_{it}^{pre}}, G_{it}^{pre}\right) - \text{Cov}\left(\frac{G_{it}^{post}}{G_{it}^{pre}}, G_{it}^{pre}\right) = E(G_{it}^{pre}) - E(G_{it}^{post}) + C \quad (7)$$

where C involves products of the sample averages of the pre and post Ginis. Thus, the object of interest is

$$\beta^* = \frac{E(G_{it}^{pre}) - E(G_{it}^{post}) + C}{\text{Var}(G_{it}^{pre})} \quad (8)$$

Direct estimation of this from equation 2 is problematic because the dependent variable involves the inverse of the main right-hand-side variable. We propose, instead, that the importance of market income inequality for relative redistribution be calculated by

$$\beta^* = \frac{1 - b}{\overline{G^{pre}}}, \quad (9)$$

where $\overline{G^{pre}}$ is the sample average of market income inequality and b is the linear projection coefficient discussed above.

5 Results

5.1 Determinants of redistribution

We first present the regression results which follow the conventional analysis of the determinants of redistribution. That is, we explain either absolute redistribution or relative redistribution by market income inequality and covariates. Absolute redistribution is measured as difference between market and disposable income Ginis. As there are only few market income Gini observations for middle- and lower-income countries, we use the gross income Gini as a proxy for the market income Gini. Relative redistribution is defined absolute redistribution divided by market income gini. In constructing the variable, we again use gross income Gini data for middle- and lower-income countries. One of the key interests is to examine the impact of underlying gross inequality on redistribution.

Table 1 presents these results. The dependent variable in Columns 1 and 2 is absolute redistribution; relative redistribution is examined in Columns 3 and 4. All models include time period dummies, and Columns 2 and 4 also include country fixed effects. The models without country fixed effects suggest that income level is, as expected, closely positively linked to redistribution. Also countries

Table 1 Determinants of redistribution, full sample.

VARIABLES	(1) Absol. redist	(2) Absol. redist	(3) Rel. redist	(4) Rel. redist
Log GDP per capita	2.276*** (0.485)	-1.272 (2.067)	4.990*** (0.916)	-5.800 (4.431)
Log population	-0.944** (0.397)	1.430 (5.361)	-2.159*** (0.751)	2.834 (9.719)
Openness	1.043 (1.498)	-0.416 (2.668)	1.466 (2.975)	0.749 (5.391)
Gross income Gini	0.342*** (0.0748)	0.677*** (0.0862)	0.232* (0.123)	0.841*** (0.151)
Democracy	-0.351* (0.199)	-0.351** (0.165)	-0.570* (0.329)	-0.674** (0.296)
Federation	0.344 (1.087)		0.197 (2.092)	
Ethnic fractionalization	-6.184*** (2.311)		-10.80** (4.323)	
Commodity exporter	-4.615*** (1.629)		-9.017*** (2.938)	
Plurality system	-1.681* (0.895)		-4.674*** (1.643)	
Observations	155	170	155	170
R-squared	0.581	0.548	0.631	0.421
Time dummies	X	X	X	X
Country dummies		X		X
Number of countries		65		65

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

with a greater population tend to redistribute less. In line with previous findings, countries which are ethnically divided have a lower desire to redistribute⁶ Countries with plural electoral system or commodity exports redistribute less. Somewhat unexpectedly, more democratic also appear to have lower redistribution. With the fixed effect models, most of the political and institutional variables are dropped out due to lack of within-country variation. With the fixed effects model, income level or population is no longer a statistically significant determinant of redistribution.

Underlying or inherent inequality is positively linked with greater redistribution. This finding is, loosely speaking, in line with political economy models (such as the median voter theorem), although one needs to bear in mind the caveats raised by Milanovic (2000), discussed in Section 2. This observation is also in line with the optimal tax tradition, initiated by Mirrlees (1971). There, redistribution at the optimum is increasing with increased pre-tax inequality.

⁶We also ran models without African countries to see whether the large extent of ethnic heterogeneity there explains this result. However, the negative impact of ethnic heterogeneity also holds for the sample without the African countries.

These results should be interpreted cautiously because of potential endogeneity concerns. It is certainly possible that redistribution also affects some of the right-hand side variables. For example, if the efficiency-equity trade off were to hold, the level of GDP per capita would be dependent on the extent of redistribution. To address these concerns, we also run models where all the right-hand side variables are lagged by one period. This is, admittedly, partly problematic as it further reduces the sample size.

The result from using the lagged values are presented in Table 2. For comparability reasons, we also present results, in Table 3, from the previous regression (with contemporaneous right-hand side variables), but with the same sample as in Table 2. In these regressions, ethnic heterogeneity, and to some extent democracy, are no longer positively linked with redistribution, as was the case for the full sample. However, as this observation is true regarding the results in both Tables 2 and 3, it seems to be driven by the data availability rather than possible endogeneity concerns. In the results of regressions without country dummies in both Table 2 and 3, the positive impact of income and the negative impact of commodity exporting and greater majoritarianism remain valid.⁷ All in all, this further analysis tends to suggest that endogeneity with respect to other determinants apart from the market income Gini is unlikely to be a major concern in these analyses.

Finally, in Table 4 we report the results from a regression where instead of redistribution, the dependent variable is our concept of net Gini (net income or consumption). Column 1 reports results from pooled OLS and Column 2 from country fixed effects regression. The results suggest that net income inequality is driven to a high extent by gross income inequality. The coefficient for the gross income inequality, b can be used to derive the implied coefficient on redistribution, as shown in the previous section. This term, which is $1 - b$, is also reported in the Table (see 'Implied abs beta') and is reasonably close to the results reported for the gross income Gini term in Table 1. This means that in this case, perhaps by chance, the mechanical correlation present in the redistribution equation did not cause a large bias. However, the bias appears to be much larger in the implied elasticity of relative beta. This can be seen by comparing the 'Implied rel beta' terms in Table 4 with the coefficient of Gross income Gini in Columns 3 and 4 of Table 1.

A concern when using data on Gini coefficients from many different sources is that the incomes on which the Gini are estimated use several different equivalence scales to account for household economies of scale. In our sample, roughly in one third of the cases income use a per capita scale, an actual household adult equivalence scale, or use no adjustment, respectively. It is plausible this affects the Gini estimates. To control for this source of variation in the Ginis, we estimate the fixed effects models of Tables 1 and 4 with data in which the same equivalence scale is used within each country. This reduces sample sizes by ten to fifteen percent, which reflects the fact that most of the full data has already set the equivalence scale for each country. The results, which are available on request, do not substantially change.

As a further robustness check, we also estimate the OLS models of Tables 1 and 4 including continent dummies to control for regional differences in redistributive preferences. The results, which are available upon request, show that the coefficient estimates on population, ethnic fractionalization and plurality are cease to be statistically significant. The gross income Gini, however, still increases redistribution in a statistically significant manner.

⁷In the model with lagged variables and country fixed effects, many of the coefficients change signs. This is understandable in a model which is identified from time differences within a country if the series have mean reversion or other related time series properties.

Table 2 Determinants of redistribution, lagged right-hand side variables.

VARIABLES	(1) Absol. redistrib	(2) Absol. redistrib	(3) Rel. redistrib	(4) Rel. redistrib
Log GDP per capita, lagged	1.829*** (0.553)	-3.189 (6.237)	4.054*** (1.045)	-8.181 (11.62)
Log population, lagged	-0.887** (0.441)	-22.91** (9.190)	-2.212** (0.873)	-32.11** (15.67)
Openness, lagged	-1.937 (1.215)	-7.062 (5.148)	-4.164* (2.392)	-13.29 (10.26)
Gross income Gini, lagged	0.241*** (0.0739)	-0.162 (0.208)	0.143 (0.139)	-0.402 (0.375)
Democracy, lagged	0.0794 (0.178)	3.170*** (0.965)	0.276 (0.351)	5.072*** (1.762)
Federation, lagged	0.568 (1.115)		0.178 (2.253)	
Ethnic fractionalization, lagged	-1.259 (2.586)		-0.901 (5.233)	
Commodity exporter, lagged	-3.298*** (0.713)		-5.192*** (1.599)	
Plurality system, lagged	-5.294*** (1.252)		-11.10*** (2.472)	
Observations	76	85	76	85
R-squared	0.771	0.613	0.804	0.572
Time dummies	X	X	X	X
Country dummies		X		X
Number of countries		41		41

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3 Determinants of redistribution, using the sample in the analysis with lagged variables.

VARIABLES	(1) Absol. redistrib	(2) Absol. redistrib	(3) Rel. redistrib	(4) Rel. redistrib
Log GDP per capita	2.615*** (0.548)	-11.05** (5.102)	5.366*** (1.125)	-17.72** (8.522)
Log population	-0.774 (0.469)	-0.537 (6.946)	-2.220** (0.983)	1.330 (13.06)
Openness	-0.446 (1.179)	-1.559 (3.600)	-1.638 (2.508)	-3.751 (6.152)
Gross income Gini	0.394*** (0.0671)	0.478** (0.185)	0.283* (0.153)	0.344 (0.342)
Democracy	-0.128 (0.161)	-0.958 (0.617)	-0.179 (0.340)	-2.077* (1.055)
Federation	0.179 (1.232)		0.571 (2.543)	
Ethnic fractionalization	-1.554 (2.806)		-3.694 (5.846)	
Commodity exporter	-3.058*** (0.976)		-4.950** (2.240)	
Plurality system	-4.723*** (1.004)		-9.712*** (2.200)	
Observations	75	85	75	85
R-squared	0.793	0.647	0.797	0.561
Time dummies	X	X	X	X
Country dummies		X		X
Number of countries		41		41

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4 Determinants of net inequality.

VARIABLES	(1) Net income Gini	(2) Net income Gini
Log GDP per capita	-3.344*** (0.600)	-5.967*** (1.655)
Log population	-0.239 (0.449)	5.503 (3.648)
Openness	-3.180** (1.352)	2.948 (1.979)
Gross income Gini	0.752*** (0.0952)	0.252*** (0.0918)
Democracy	0.309 (0.340)	0.336 (0.839)
Federation	1.461 (1.215)	
Ethnic fractionalization	4.141* (2.260)	
Commodity exporter	5.695*** (1.019)	
Plurality system	2.476*** (0.904)	
Observations	127	131
R-squared	0.801	0.594
Time dummies	X	X
Country dummies		X
Implied abs beta	0.248	0.748
Implied rel beta	0.541	1.632
Number of countries		49

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

5.2 Inequality, redistribution and growth

Next we replicate the analysis of Ostry, Berg, and Tsangarides (2014) with WIID data. They study determinants of medium-term aggregate growth, and their main finding is that income inequality decreases growth whereas redistribution does not affect the growth. For income inequality and redistribution they use SWIID data that contains a substantial number of imputed Gini observations. We examine whether the main results of their study prevail with WIID data.⁸

Table 5 presents the results of a model in which five-year average redistribution is regressed on five-year averages of the net income Ginis, the natural logarithm of GDP per capita, and time period dummies (Ostry, Berg, and Tsangarides 2014, Table 2). Redistribution is defined, as above, as the difference between market and net income Gini. The model is estimated with country fixed effects; standard errors are robust with respect to within-country autocorrelation in the error term.⁹ Table 5 is arranged so that the first results column represents results using the Ostry, Berg, and Tsangarides (2014) data (i.e. the Gini coefficients are from SWIID), the third column using WIID data, and column in between with use the Ostry, Berg, and Tsangarides (2014) data, but limit the sample to those observations that are included in the WIID sample.

As can be seen, net income inequality is associated with a higher level of absolute redistribution, as noted in the previous section. However, the caveat for endogeneity of market income inequality applies. Sample size decreases to one fourth as we move from SWIID to WIID data. Sample size for the second and third column do not equalize because SWIID data Ostry, Berg, and Tsangarides (2014) use do not have all the Gini observations there exist in WIID. Consequently, differences in results of the second and third column are due to that and the method Solt (2013) uses when estimating income inequality in SWIID.

Table 5 Replication of Ostry et al. (2014) Table 2. Determinants of redistribution, full sample.

	Ostry et. al	Ostry et al. with WIID observations	WIID
Market inequality	0.483*** (0.0523)	0.629*** (0.1010)	0.681*** (0.0646)
Log of GDP per capita	1.469 (0.9377)	4.083** (1.5987)	-1.862 (2.1682)
Constant	-28.392*** (8.1707)	-56.002*** (15.8095)	-3.180 (20.6917)
Observations	829	187	206
R-squared	0.8797	0.9768	0.9472

Standard errors in parentheses

FE estimation with five-year period dummies.

Clustered Standard errors robust for intragroup (country) correlation.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The main focus of the Ostry, Berg, and Tsangarides (2014) study is whether inequality and redistribution affect economic growth. They consider both inequality and redistribution as, for example,

⁸For other variables we use their data and empirical specifications unless stated otherwise. Their data and Stata do-files for the estimations can be downloaded on <https://www.imf.org/external/pubs/cat/longres.aspx?sk=41291.0>.

⁹For brevity, we only report results for the full sample. Results for OECD and non-OECD subsamples are available on request. The OECD subsample results are inline with the full sample results. With the non-OECD subsample and SWIID data, income level affects redistribution with a positive point estimate.

inequality may not impede growth while redistributive efforts do. They regress five-years average real GDP growth rate on market income inequality, redistribution, and other controls. The models are estimated for non-overlapping five-year periods starting from 1960. Investment is defined as investment share of GDP (in percentages), total education average years of primary and secondary schooling, political institutions Polity IV score, openness, i.e. the sum of exports and imports per GDP, and debt liabilities external debt per GDP. The terms of trade shock variable takes the value 1 when the change in the terms of trade is among three bottom decile groups. To address likely endogeneity of the right hand side variables, they estimate the models with System GMM. They use Windmeijer's small-sample correction for the point estimate variances.

The results are shown in Table 6. Ostry, Berg, and Tsangarides (2014) estimate four models with differing number of regressors. To be able to assess extent to which differences in the estimation results are due to different data and different sample composition, the table is organized for each model as above.

As can be seen Ostry, Berg, and Tsangarides (2014) find that net income inequality is detrimental for growth whereas redistribution has not statistically significant effect the on growth. Other covariates affect on the growth in expected manner.¹⁰

¹⁰We cannot exactly replicate results reported in their paper. However, this is probably due to fact the data set on their website is more up-to-date than the one used for the estimations in their paper.

Table 6 Replication of Ostry et al. (2014) Table 3. Regression of five-year economic growth.

	Baseline (BL)	BL with WIID observations	BL, WIID	BL + ctrls (2)	BL + ctrls (2) with WIID obs.	BL + ctrls (2), WIID
Log(initial income)	-0.00701** (-2.00)	-0.00775 (-1.37)	-0.00632 (-1.35)	-0.00822** (-2.30)	-0.0180*** (-3.78)	-0.00563 (-1.05)
Net inequality	-0.00144*** (-3.19)	-0.00190** (-2.38)	-0.00152** (-2.19)	-0.000892*** (-2.63)	-0.000905 (-1.10)	-0.000888 (-1.25)
Redistribution	0.0000413 (0.08)	-0.00153** (-2.03)	-0.00125* (-1.80)	0.000303 (0.54)	-0.000401 (-0.49)	-0.00232** (-2.23)
Log(investment)				0.0245*** (3.13)	0.0337*** (2.73)	0.000993 (0.05)
Log(population growth)				-0.0158 (-0.84)	-0.0817** (-2.33)	-0.0947** (-2.12)
Log(total education)						
Large negative terms of trade shock						
Political institutions						
Openness						
Debt liabilities						
Constant	0.137*** (3.32)	0.183*** (2.64)	0.155*** (2.97)	0.0806* (1.72)	0.271*** (4.11)	0.305*** (3.15)
Observations	828	153	153	828	153	153
Number of countries	130	52	52	130	52	52
Sargan	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hansen	0.3805	0.9998	0.9989	0.3802	1.0000	1.0000
AR1	0.0000	0.1495	0.0797	0.0000	0.0992	0.0580
AR2	0.1641	0.6879	0.6888	0.1187	0.9014	0.8582
Number of instruments	117	92	77	133	98	85
longtable						

t statistics in parentheses

System GMM estimation with time period dummies and Windmeijer's finite-sample corrected standard errors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6 Ostry et al. (2014) Table 3 continued.

	BL + ctrls (3)	BL + ctrls (3) with WIID obs.	BL + ctrls (3), WIID	BL + ctrls (4)	BL + ctrls (4) with WIID obs.	BL + ctrls (4), WIID
Log(initial income)	-0.0143*** (-3.79)	-0.0143*** (-2.17)	-0.00918 (-1.10)	-0.0125** (-2.51)	-0.0237 (-1.10)	-0.0258* (-2.54)
Net inequality	-0.000707*** (-2.64)	-0.00115 (-1.44)	-0.000301 (-0.41)	-0.00106* (-1.95)	-0.00128 (-0.78)	-0.000625 (-0.58)
Redistribution	0.000187 (0.47)	-0.000786 (-1.09)	-0.00122 (-1.50)	0.000270 (0.51)	-0.000767 (-0.96)	-0.000808 (-0.66)
Log(investment)	0.0231** (2.56)	0.0278** (2.16)	0.0166 (0.90)	0.00691 (0.45)	0.0413 (1.00)	0.0235 (0.93)
Log(population growth)	-0.0232 (-1.19)	-0.0451 (-1.54)	-0.0904** (-2.34)	-0.00564 (-0.30)	-0.0123 (-0.22)	-0.0493 (-1.03)
Log(total education)	0.0219*** (2.85)	0.0178 (1.59)	0.0306 (0.81)	0.0152 (1.48)	0.0404 (1.21)	0.0368 (1.15)
Large negative terms of trade shock				-0.0472*** (-3.02)	-0.0314 (-1.20)	-0.0303 (-1.25)
Political institutions				-0.00109 (-1.43)	0.00213 (0.97)	0.00306 (0.55)
Openness				0.0133 (1.49)	-0.0112 (-0.26)	-0.000927 (-0.04)
Debt liabilities				-0.0176*** (-2.79)	0.000763 (0.06)	-0.00425 (-0.77)
Constant	0.102*** (2.58)	0.161*** (2.77)	0.175 (1.40)	0.156** (2.44)	0.108 (0.42)	0.229 (1.33)
Observations	751	150	150	558	116	116
Number of countries	110	50	50	79	37	37
Sargan	0.0000	0.0000	0.0002	0.0027	0.0001	0.0016
Hansen	0.9502	1.0000	0.9998	0.8605	0.9999	1.0000
AR1	0.0000	0.1756	0.0674	0.0000	0.1346	0.1096
AR2	0.1704	0.9387	0.3560	0.5418	0.4313	0.2326
Number of instruments	139	101	90	100	76	67

t statistics in parentheses

System GMM estimation with time period dummies and Windmeijer's finite-sample corrected standard errors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Estimations with WIID data lend support for their first model. However, models with increasing number of control variables, neither income inequality nor redistribution has anymore statistically significant effect on the growth. This may, however, be due to sample composition. Sample size drops, once again, drastically when we move from SWIID to WIID sample. As actual income inequality data is mainly available for developed countries, this tilts the sample composition towards the country group; around three fourths of the SWIID sample for the baseline model are from non-OECD countries, while only one third of the observations with WIID data are from non-OECD countries. Table 7 reports sample composition for Ostry, Berg, and Tsangarides (2014) Table 6 in which they estimate baseline model of Table 6 of this paper. Different SWIID samples are based on different inclusion criteria for different SWIID sample¹¹.

¹¹See footnote for Table 4 in Ostry, Berg, and Tsangarides (2014).

Table 7 Number of five-years average redistribution observations with different Ostry et al. (2014) samples and WIID.

Income group	SWIID, full sample	SWIID, baseline sample	SWIID, restricted sample	WIID
Low income	137	107	31	20
Lower middle income	233	149	82	59
Upper middle income	300	233	110	8
High income: nonOECD	94	91	52	3
High income: OECD	285	285	207	127
Total	1049	865	482	217

To explore more thoroughly this, we estimate models of Table 6 with SWIID data separately for OECD and non-OECD countries. Results are shown in Tables 8 and 9. They confirm the suspicion. In OECD countries, neither net income inequality nor redistribution (except in the Baseline model) have a statistically significant effect on the growth. The results for the non-OECD countries, however, are more in line with Ostry, Berg, and Tsangarides (2014) results; income inequality is detrimental for the growth but redistribution has not statistically significant effect. So, the conclusion they make is subject to country sample.

Table 8 Ostry et al. (2014) Table 3 with only OECD countries. (The effect of inequality and redistribution on growth)

	BL	BL + ctrls (1)	BL + ctrls (2)	BL + ctrls (3)
Log(initial income)	-0.136*** (-3.26)	-0.0652 (-1.29)	-0.0551*** (-2.84)	-0.0976 (-0.63)
Net inequality	0.000418 (0.30)	-0.00198 (-0.46)	0.00131 (0.47)	-0.00770 (-0.23)
Redistribution	0.00343** (2.04)	0.0000574 (0.05)	0.000381 (0.45)	-0.00192 (-0.13)
Log(investment)		0.0480 (0.85)	0.0402 (1.15)	0.0266 (0.09)
Log(population growth)		-0.0229 (-0.30)	-0.0692 (-0.74)	0.216 (0.28)
Log(total education)			-0.0475* (-1.73)	0.0331 (0.36)
Large negative terms of trade shock				-0.0849 (-0.20)
Political institutions				0.00774 (1.35)
Openness				0.291 (0.23)
Debt liabilities				0.0233 (0.10)
Constant	1.357*** (3.33)	0.638 (1.08)	0.653*** (2.68)	0.476 (0.31)
Observations	220	220	220	182
Number of countries	24	24	24	22
Sargan	0.1446	0.0003	0.0000	0.0807
Hansen	1.0000	1.0000	1.0000	1.0000
AR1	0.4037	0.0343	0.0072	0.9549
AR2	0.2388	0.1396	0.2314	0.8369
Number of instruments	117	132	137	100

t statistics in parentheses

System GMM estimation with time period dummies and Windmeijer's finite-sample corrected standard errors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9 Ostry et al. (2014) Table 3 with only non-OECD countries. (The effect of inequality and redistribution on growth)

	BL	BL + ctrls (1)	BL + ctrls (2)	BL + ctrls (3)
Log(initial income)	-0.000259 (-0.07)	-0.00596 (-1.39)	-0.0107** (-1.98)	-0.00946 (-0.98)
Net inequality	-0.00115*** (-2.68)	-0.000651* (-1.75)	-0.000628** (-2.01)	-0.000918 (-1.64)
Redistribution	0.000681 (1.10)	0.000895 (1.64)	0.000259 (0.52)	0.000696 (0.89)
Log(investment)		0.0220*** (2.59)	0.0244** (2.40)	0.0167 (0.85)
Log(population growth)		-0.00780 (-0.54)	-0.0102 (-0.65)	0.00463 (0.25)
Log(total education)			0.0223** (2.29)	0.0193* (1.65)
Large negative terms of trade shock				-0.0431** (-2.41)
Political institutions				-0.00167* (-1.65)
Openness				-0.000406 (-0.03)
Debt liabilities				-0.0142 (-0.94)
Constant	0.0671** (2.05)	0.0409 (0.95)	0.0360 (0.96)	0.0723 (0.64)
Observations	608	608	531	376
Number of countries	106	106	86	57
Sargan	0.0000	0.0000	0.0000	0.0597
Hansen	0.7280	0.9230	0.9997	0.9995
AR1	0.0000	0.0000	0.0000	0.0000
AR2	0.0690	0.0588	0.1230	0.3834
Number of instruments	117	132	137	100

t statistics in parentheses

System GMM estimation with time period dummies and Windmeijer's finite-sample corrected standard errors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ostry, Berg, and Tsangarides (2014) use net income inequality as one of the growth determinants in their regression when you could think market income inequality better reflects underlying inequality. To check if the conclusions prevail with market income inequality data, we estimate models of Table 6 with market income inequality instead of net income inequality as regressor. Results with market income inequality data are very much in line with the results with net income inequality data ¹². As another variant we drop redistribution as a regressor to save observations with WIID data, which more than doubles the sample size. Results are in line with Table 6. With SWIID sample unit increase in Gini coefficient is associated with 0.05 to 0.15 percentage points decrease in annual economic growth. However with WIID sample we find no statistically significant evidence of detrimental effect of income inequality on the growth.

Lastly we assess robustness of Ostry, Berg, and Tsangarides (2014) results on model specification. As can be seen in Table 6, The Sargan test of over-identifying restrictions suggests that not all GMM-style instrument variables are exogenous as assumed. Also, some of the Hansen test p-values are implausible good. According to Roodman (2009), large instrument sets, which System GMM estimators are susceptible to, weaken the Hansen test.

The Arellano-Bond autocorrelation test suggest no autocorrelation in idiosyncratic disturbance. However, we check if using longer lags for the instrumental variables would help with the exogeneity. Based on the Sargan statistics it does not. Estimation results with the model yield less robust evidence for statistically significant effect of income inequality on the growth.

Number of instrumental variables in System GMM increases fast with panel length, especially if several lags are used as instrumental variable for given variable and time period. To limit this, we estimate model of Table 6 with SWIID data and only using first relevant lag for the instrument variables. The estimation results with the specification are very much in line with the original ones. As a final specification test we further simplify the estimated model by omitting possible endogeneity of the regressors and estimating the model with fixed effects estimation. Results are similar with Table 6 except that income inequality is no more statistically significant.

6 Conclusion

The paper had three main goals. First, we used the WIID data to describe the availability and the extent of redistribution in all countries in the world. Since one needs to know both before and after government intervention inequality levels, whereas for many countries only disposable income inequality or consumption inequality figures are available, information about inequality is surprisingly and regrettably sparse. This is the case especially for developing countries.

Second, we contributed to the empirical literature on the determinants of redistribution methodologically. Earlier, influential, analysis in the field has not paid attention to the problem that when inequality (the difference between market and net income inequality) is explained by the underlying market inequality, the regression is plagued by mechanical correlation between the left-hand side and right-hand side variables. We suggested a way to overcome this bias by a procedure that uses the coefficient from a regression of net inequality on market inequality to deduce the association between redistribution and market income inequality.

Third, we examined the robustness of recent work by Ostry, Berg, and Tsangarides (2014) who use the SWIID data to examine the impacts of redistribution on growth. The study may potentially suffer from weaknesses in the redistribution data, as it is to a large extent based on imputations. Using the

¹²For brevity, we do not report this and rest of the results here. They are, however, available upon request.

arguably preferred WIID data set, which consists of real rather than imputed observations, and their empirical techniques implies that some of the results in their analysis change. Net income inequality does not affect growth negatively, but the insignificant relationship between redistribution and growth (in a model with control variables) remains. The latter finding is important in that the lack of statistical significance in the original Ostry, Berg, and Tsangarides (2014) analysis could have been a result of an attenuation bias created by measurement error. The findings of this paper can be interpreted as lending support to the ideas that there would not be an efficiency-equity tradeoff in a cross country setting, but one also needs to bear in mind that this result is based on a limited number of observations.

References

- Alesina, Alberto and Roberto Perotti (1996). "Income distribution, political instability, and investment". In: *European Economic Review* 40.6, pp. 1203–28.
- Alesina, Alberto and Dani Rodrik (1994). "Distributive Politics and Economic Growth". In: *The Quarterly Journal of Economics* 109.2, pp. 465–90. ISSN: 00335533, 15314650.
- Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg (2003). *Fractionalization*. Working Paper 9411. Cambridge, MA: National Bureau of Economic Research. DOI: [10.3386/w9411](https://doi.org/10.3386/w9411). URL: <http://www.nber.org/papers/w9411>.
- Atkinson, Anthony B and Andrea Brandolini (2001). "Promise and Pitfalls in the Use of Secondary Data-Sets: Income Inequality in OECD Countries as a Case Study". In: *Journal of Economic Literature* 39.3, pp. 771–99.
- Barro, Robert J (2000). "Inequality and Growth in a Panel of Countries". In: *Journal of Economic Growth* 5.1, pp. 5–32.
- Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh (2001). "New tools in comparative political economy: The Database of Political Institutions". In: *World Bank Economic Review* 15.1, pp. 165–76.
- Bertola, Giuseppe (1993). "Factor Shares and Savings in Endogenous Growth". In: *American Economic Review* 83.5, pp. 1184–98.
- Economist (2014). "Free exchange: Inequality v growth". In: *The Economist* 1st March 2014.
- Ferreira, Francisco H, Nora Lustig, and Daniel Teles (2015). "Appraising cross-national income inequality databases: An introduction". In: *The Journal of Economic Inequality* 13.4, pp. 497–526. ISSN: 1569-1721. DOI: [10.1007/s10888-015-9316-0](https://doi.org/10.1007/s10888-015-9316-0). URL: <http://dx.doi.org/10.1007/s10888-015-9316-0>.
- Garcia-Penalosa, Cecilia, Eve Caroli, and Philippe Aghion (1999). "Inequality and Economic Growth: The Perspective of the New Growth Theories". In: *Journal of Economic Literature* 37.4, pp. 1615–60.
- Guardian (2014). "IMF study finds inequality is damaging to economic growth". In: *The Guardian* 26th February.
- Jenkins, Stephen P. (2015a). "World income inequality databases: an assessment of WIID and SWIID". In: *The Journal of Economic Inequality* 13.4, pp. 629–671.
- Jenkins, Stephen P (2015b). "World income inequality databases: an assessment of WIID and SWIID". English. In: *The Journal of Economic Inequality* 13.4, pp. 629–71. ISSN: 1569-1721. DOI: [10.1007/s10888-015-9305-3](https://doi.org/10.1007/s10888-015-9305-3). URL: <http://dx.doi.org/10.1007/s10888-015-9305-3>.
- Karabarbounis, Loukas (2011). "One Dollar, One Vote". In: *Economic Journal* 121.553, pp. 621–51.
- Lane, Philip R and Gian Maria Milesi-Ferretti (2007). "The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970-2004". In: *Journal of International Economics* 73.2, pp. 223–50.
- Luebker, Malte (2014). "Income Inequality, Redistribution, and Poverty: Contrasting Rational Choice and Behavioral Perspectives". In: *Review of Income and Wealth* 60.1, pp. 133–54.
- Meltzer, Allan H and Scott F Richard (1981). "A Rational Theory of the Size of Government". In: *Journal of Political Economy* 89.5, pp. 914–27.
- Milanovic, Branko (2000). "The median-voter hypothesis, income inequality, and income redistribution: an empirical test with the required data". In: *European Journal of Political Economy* 16.3, pp. 367–410.

- Milanovic, Branko (2010). “Four critiques of the redistribution hypothesis: An assessment”. In: *European Journal of Political Economy* 26.1, pp. 147–54.
- Mirrlees, James A (1971). “An exploration in the theory of optimal income taxation”. In: *Review of Economic Studies* 38, pp. 175–208.
- Okun, Arthur M (1975). *Equality and efficiency. The big tradeoff*. Washington D.C.: Brookings Institution.
- Ostry, Jonathan David, Andrew Berg, and Charalambos G. Tsangarides (2014). *Redistribution, Inequality, and Growth*. IMF Staff Discussion Notes 14/02. International Monetary Fund.
- Perotti, Roberto (1993). “Political Equilibrium, Income Distribution, and Growth”. In: *Review of Economic Studies* 60.4, pp. 755–76.
- (1996). “Growth, Income Distribution and Democracy”. In: *Journal of Economic Growth* 1.2, pp. 149–87.
- Persson, Torsten and Guido Tabellini (1994). “Is Inequality Harmful for Growth?” In: *American Economic Review* 84.3, pp. 600–21.
- Rodrik, Dani (1999). “Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses”. In: *Journal of Economic Growth* 4.4, pp. 385–412.
- Roodman, David (2009). “A Note on the Theme of Too Many Instruments”. In: *Oxford Bulletin of Economics and Statistics* 71.1, pp. 135–58.
- Scervini, Francesco (2012). “Empirics of the median voter: democracy, redistribution and the role of the middle class”. In: *Journal of Economic Inequality* 10.4, pp. 529–50.
- Solt, Frederick (2009). “Standardizing the World Income Inequality Database”. In: *Social Science Quarterly* 90.2, pp. 231–42.
- (2013). *SWIID. The Standardized World Income Inequality Database, version 4.0*. Statistical database. URL: <http://myweb.uiowa.edu/fsolt/swiid/swiid.html>.
- Tuomala, Matti and Hannu Tanninen (2005). “Inherent Inequality and the Extent of Redistribution in OECD Countries”. In: *CESifo DICE Report* 3.1, pp. 48–53.
- UNU-WIDER (2015). *World Income Inequality Database (WIID3.3)*. Statistical database, <https://www.wider.unu.edu/download/WIID3.3>.
- World Bank (2016). *World Development Indicators*. URL: <http://data.worldbank.org/data-catalog/world-development-indicators..>

Appendix A Data sources and definitions

A.1 Data Definitions and Sources for Redistribution Regressions

Variable	Definition	Source
Absolute redistribution	Difference between market and disposable income Ginis. If market income Gini is unavailable gross income Gini is used.	World Income Inequality Database (WIID) version 3.3
Relative redistribution	Relative difference between market and disposable income Ginis.	WIID version 3.3
Log GPD per capita		World Development Indicators (WDI)
Log population		WDI
Openness	Imports + exports of GDP	WDI
Gross income Gini	Gross income Gini.	WIID 3.3
Ethnic fractionalization	Index of ethnic fractionalization (involves a combination of racial and linguistic characteristics). Higher the index number, higher the fractionalization.	Alesina et al. (2003)
Democracy	Degree of autocracy and democracy of government. Taking increment values between -10 and 10 with the minimum referring to very autocratic and the maximum very democratic government.	Polity IV dataset.
Federation	Dummy variable 1 for federal governing system	IMF
Plurality system	Electoral rule for the majority House seats. One for plurality rule, zero for proportionality rule.	Database of Political Institutions Beck et al. (2001)
Commodity exporter	Dummy variable taking value one if average net exports per GDP exceed ten percent for agricultural products, or fuels, or ores over 1990-2009.	World Bank WDI and authors' calculation

A.2 Data Definitions and Sources for Ostry, Berg, and Tsangarides (2014) Replication

All the variables are five-years averages, unless noted otherwise. For SWIID and control variables we use the same data as Ostry, Berg, and Tsangarides (2014).¹³ For WIID, we use the latest version 3.3.

¹³Available at <http://www.imf.org/external/pubs/ft/sdn/2014/data/sdn1402.zip>

Variable	Definition	Source
Log(Initial income)	Logarithm of real GDP in the beginning of each five years period.	Penn World Tables (PWT) 7.1
Net inequality	Net income Gini	SWIID 3.1 and WIID 3.3.
Redistribution	Difference of market and net income Gini	SWIID 3.1 and WIID 3.3.
Log(investment)	Log of investment per GDP in percentage	PWT 7.1
Log(population growth)	Log of five-years relative population change plus five.	PWT 7.1
Log(total education)	Average years of primary and secondary school education in logarithm.	PWT 7.1
Large negative terms of trade shock	Dummy 1 when terms of trade is among bottom three deciles of countries.	IMF WEO
Political institutions	Variable Polity2 in Polity IV database.	Polity IV dataset.
Openness	Share of exports plus imports per GDP.	PWT 7.1
Debt liabilities	External debt per GDP.	Lane and Milesi-Ferretti (2007)

Appendix B WIID data set manipulation

As Ostry, Berg, and Tsangarides (2014), we use five-year period averages inequality and redistribution measures. The welfare definition of the Gini index describes if the data is for market income, net income, or consumption Gini. There are several categories in the welfare definition. We define the Gini observation to describe net income Gini if welfare definition in WIID takes one of the values

- "Monetary Income, Disposable"
- "Earnings, Net"
- "Monetary Income, Disposable (excluding property income)"
- "Monetary Income, Disposable, excl. self-empl. and property income",

On the other hand, for consumption Gini, welfare definition is one of

- "Income/Consumption"
- "Consumption"
- "Income/Consumption" and variable Region takes value "Americas".

Market income Gini has welfare definition of one of

- "Factor Income"
- "Taxable Income, property income excluded"
- "Taxable Income, Gross incl deductions"
- "Taxable Income, Gross"
- "Earnings, Gross"
- "Income, Gross"
- "Monetary Income, Gross"
- "Market Income",

In WIID for some country-year observations there are multiple observations for any of the Gini measures. Instead of simply averaging over the duplicate observation, we prefer picking observation we think has the best quality information. For that we use equivalence scale, age coverage, and area coverage associated with Gini observation. The following procedure is used to pick the observation

1. Equivalence scale (variable Equivsc_new): First if the observation takes value "Household Adult Equiv", then rest,
2. Age coverage (AgeCovr_new): First "All", then rest,
3. Area coverage (AreaCovr_new): First "All", then rest.

If still even, then have average of those observations.

In the analysis for each country-year observation we combine net income and consumption income Gini as net income Gini. For that we first use net income Gini information. If that is missing, we use consumption Gini.

After this we have five-year period (1960-1964, 1965-1969, and so on) averages for each country.