WP/16/244

IMF Working Paper

Inequality and Growth: A Heterogeneous Approach

Francesco Grigoli, Evelio Paredes, and Gabriel Di Bella

INTERNATIONAL MONETARY FUND

 WP/16/244

IMF Working Paper

Western Hemisphere Department

Inequality and Growth: A Heterogeneous Approach

Prepared by Francesco Grigoli, Evelio Paredes, and Gabriel Di Bella

Authorized for distribution by Alfredo Cuevas

December 2016

Abstract

The combination of stagnant growth and high levels of income inequality renewed the debate about whether a more even distribution of income can spur economic activity. This paper tests for cross-country convergence in income inequality and estimates its impact on economic growth with a heterogeneous panel structural vector autoregression model, which addresses some empirical challenges plaguing the literature. We find that income inequality is converging across countries, and that its impact on economic growth is heterogeneous. In particular, while the median response of real per capita GDP growth to shocks in income inequality is negative and significant, the dispersion around the estimates is large, with at least one fourth of the countries in the sample presenting a positive effect. The results suggest that the negative effect is mainly driven by the Middle East and Central Asia and the Western Hemisphere across regions, and emerging markets across income levels. Finally, we find evidence that improved institutional frameworks can reduce the negative effect of income inequality on growth.

JEL Classification Numbers: E15, O15, O40

Keywords: Heterogeneity, Gini, income distribution, income inequality, income levels, growth, regions.

Author’s E-Mail Address: fgrigoli@imf.org; eparedes@hcacienda.gov.do; gdibella@imf.org

---

We thank, without implicating, Sebastian Acevedo, Alfredo Cuevas, Christian Ebeke, Davide Furceri, Metodij Hadzi-Vaskov, Melissa Jimenez, Fah Jirasavetakul, Alvar Kangur, Emmanouil Kitsios, Aart Kraay, Sergi Lanau, Grace Li, Mario Mansilla, Peter Pedroni, Adrian Peralta-Alva, Luca Ricci, Daniel Rodriguez-Delgado, Camile Sancak, Antonio Spilimbergo, and Marina Mendes Tavares for discussions, comments, and suggestions. In particular, we thank Carlos Goes for sharing his code and Adrian Robles for the excellent research assistance.

Contents

List of Figures

List of Tables

"This kind of inequality – a level we haven't seen since the Great Depression – hurts us all. When middle-class families can no longer afford to buy the goods and services that businesses are selling, it drags down the entire economy, from top to bottom."

— Barack Obama, *Osawatomie, KS, December 6 2011*

1 Introduction

Does income inequality slow economic growth? From the theoretical standpoint, the effect is ambiguous. On the one hand, a higher concentration of income in the hands of a few is reflected in reduced demand by a larger share of poorer individuals, which would invest less in education and health and grow a sense of social and political discontent, jeopardizing human capital and stability. Moreover, more inequality can exacerbate households' leverage to compensate for the erosion in relative income, empower the influence of the richer population on the legislative and regulatory processes, and motivate redistribution policies that are often blamed for slowing growth, especially when aggressive. On the other hand, a certain level of inequality endows the richer population with the means to start businesses, as well as creates incentives for individuals to increase their productivity and invest their saving, hence promoting economic growth.

The common wisdom for which more equitable societies should experience higher economic growth has generally been supported by the empirical literature. Indeed, there is a growing body of econometric studies that finds a negative and significant relationship between inequality and long-term growth using the Generalized Method of Moments (GMM) (Castelló-Climent, 2010; Halter et al., 2014; Ostry et al., 2014; Dabla-Norris et al., 2015). These, however, have been questioned on empirical grounds. Kraay (2015) recently pointed out that the GMM estimator is not suitable for growth regressions as internal instruments are weak, and once he corrects for it, he finds no evidence of either a positive or a negative effect of income inequality on growth. More generally, omitted variables, measurement error, reverse causality, dynamics, cross-section dependence, and homogeneity across countries all proved daunting challenges and can critically bias the inference of the results.[1]

This paper revisits the relationship between income inequality and economic growth employing a data-driven, recently developed, agnostic approach. Its contribution is threefold. First, it studies the stationarity properties of income inequality and tests for its cross-country convergence as well as for the existence of convergence clubs. Second, departing from the literature, it employs the heterogeneous Panel Structural Vector Autoregression (PSVAR) model of Pedroni (2013) to deal with some of the empirical issues raised in previous studies. Under a proper identification strategy, this methodology isolates the causal relationship between the variables of interest without resorting to internal instruments. Further, it allows for a dynamic response of growth to shocks in income inequality, with different intensities over different time horizons. By treating all the countries in the panel separately, it allows for heterogeneous slopes or responses, providing a much richer set of information about regional effects and differences across income levels. Finally, this approach helps dealing with omitted variable bias and accounts for cross-section dependence. The

[1] Barro (2000) and Forbes (2000) discuss the relevance of omitted factors. Halter et al. (2014) highlight that changes in inequality affect differently economic growth over different time horizons. Atems and Jones (2015) also look at the dynamics and point out that common factors are present. Barro (2000) and De Dominicis et al. (2006) note the importance of selecting a more homogeneous group of countries. Banerjee and Duflo (2003) provide a review of studies highlighting some of the main econometric issues.

third contribution of the paper is an investigation of the determinants of the relationship between income inequality and growth.

We find that income inequality is converging across countries, and that its impact on economic growth is heterogeneous. While countries are converging toward the sample mean, we also find that Europe, the Middle East and Central Asia, and advanced economies are experiencing within convergence. The results also suggest that the median response of real per capita GDP growth to shocks in income inequality is negative and significant for the full sample, but the dispersion around the estimates is large, with at least one fourth of the countries in the sample presenting a positive effect, implying that the focus of the literature on the average relation may be misleading. A regional analysis reveals that the negative median effect is driven by the Middle East and Central Asia and the Western Hemisphere. Across income levels, only the findings for emerging markets indicate that more inequality slows economic growth. We also find evidence suggesting that higher institutional quality, characterized by stronger control of corruption, rule of law, and government effectiveness, can reduce the negative effect of income inequality on growth.

The rest of the paper is structured as follows. Section 2 briefly reviews the empirical literature studying the effect of income inequality on growth. Section 3 presents the data used for this study and some stylized facts for the sample and across regions and income groups. Section 4 discusses the methodology, the empirical results, and their robustness. Section 5 concludes.

2 A Review of the Literature

The relationship between economic growth and income inequality has been in the research agenda of academics for a long time and has recently gained momentum. Since Kuznets (1955) argued that higher income inequality was a characteristic of poor countries and that it would eventually reverse once they develop, contradicting trends in the distribution of income across income levels revealed that much more is to be understood. Raising income inequality in advanced economies sparked a vast literature analyzing whether more unequal countries should expect less satisfactory growth performance.

There are conflicting theories trying to characterize the relationship between income inequality and growth. The ones predicting a negative relationship focus on how inequality can be harmful to sustained growth by reducing social consensus, dampening investment, and affecting education and health outcomes (Persson and Tabellini, 1994; Alesina and Perotti, 1996; Benabou, 1996; Rodrik, 1999; Easterly, 2007; Ostry and Berg, 2011). Further, more inequality is usually associated with the implementation of redistribution policies, that are often blamed for slowing economic growth (Persson and Tabellini, 1992; Alesina and Rodrik, 1991; Rajan, 2011), although Ostry et al. (2014) show that non-excessive redistribution can reduce inequality and be growth neutral. More recently, the attention shifted to look at how inequality exacerbates leverage and financial cycle creating the premises for crises (Rajan, 2011), and how the influence of the richer population on the legislative and regulatory processes can lead to financial imbalances (Stiglitz, 2012).

Other theories propose a positive relationship. These are based on the argument that inequality can rather provide incentives for innovation and higher productivity (Lazear and Rosen, 1981; Okun, 2015), foster saving and investment to the extent that rich people have a higher propensity to save (Kaldor, 1957), and endow richer individuals with the minimum capital and education needed to start some economic activity (Barro, 2000).[2] Other arguments include the one for which

[2] As noted by Forbes (2000), many of the models suggesting a negative relationship actually predict multiple

more inequality leads the median voter to elect higher taxation to finance higher capital accumulation and growth (Saint-Paul and Verdier, 1993), and the one for which inequality increases in periods of technology innovations and high growth as a result of the decline in the relative importance of initial parental-environmental conditions and concentration of abilities (Galor and Tsiddon, 1997).

The inconclusiveness of the theory's predictions casts light on the empirical research. Unfortunately, the paucity of time series for income inequality condemned the econometric studies till late 90s to exploit only the cross-section dimension of the data. These studies generally find a negative effect of income inequality on growth.[3] Since then, the increased availability of the longitudinal dimension of the data spurred a wave of research employing panel data techniques on period averages. Li and Zou (1998), Benhabib and Spiegel (2000), and Forbes (2000) employ fixed effects estimations to control for the (time-invariant) omitted country bias. The inclusion of fixed effects produced some controversial results, namely that increases in inequality promote growth.

Clearly, there is no reason to believe that the inclusion of the fixed effects fully addresses endogeneity concerns. As noted by Chudik and Pesaran (2013), common effects correlated with the variables used in the estimation are likely to be present in many panel applications. As a result, the hypothesis of cross-section independence could be violated. Thus, panel estimators such as fixed or random effects can result in misleading inference and even inconsistent estimators (Phillips and Sul, 2007). This potential source of bias, however, received virtually no attention in the empirical literature on the topic.

One of the major concerns when studying the impact of income inequality on growth is sorting the direction of causality. In the attempt to address this undetermined causality, Barro (2000) employs a three-stage least squares estimator, where instruments are lagged values of the regressors and dummies for prior colonial status. He only finds a significant effect after splitting the sample in richer and poorer countries, with the former (latter) presenting a positive (negative) effect.[4] A more recent strand of the empirical literature employs the System GMM (S-GMM) estimator, which identifies the effect of inequality on growth by instrumenting the endogenous variables with a set of lagged levels and differences of the regressors.[5] Castelló-Climent (2010), Halter et al. (2014), Ostry et al. (2014), and Dabla-Norris et al. (2015) all use S-GMM and find a negative impact of income inequality on growth. Kraay (2015) discusses the suitability of the S-GMM to growth regressions. In particular, he applies weak instrument diagnostics to the benchmark specification run by Castelló-Climent (2010), Halter et al. (2014), Ostry et al. (2014), and Dabla-Norris et al. (2015).[6] He finds pervasive evidence of weak internal instruments and, once the inference is corrected for that, he argues that it is not possible to draw conclusions about either a negative or a positive effect of inequality on growth.

Adopting a somewhat innovative approach in this literature, Atems and Jones (2015) examine both the effects of income inequality on growth and the effects of growth on income inequality for a panel of states in the United States. They employ a bivariate Vector Autoregression (VAR) model with the growth rate of real per capita income and the growth rate of the Gini, and identify the shocks by assuming that changes in income inequality affect growth with a one-year delay.

equilibria for which, under some conditions, there exists a positive relationship of income inequality with economic growth.

[3] See Benabou (2000) for a list of studies performing cross-section ordinary least squares (OLS) estimations.

[4] Contrary to other studies, Barro (2000) does not include fixed effects, arguing that they would remove relevant cross-country information.

[5] The method allows for including external instruments too, but this option has rarely been used owing to the difficulty in finding appropriate instruments. Easterly (2007) is an exception, as he uses agricultural endowments to instrument inequality and finds a negative cross-section relationship between inequality and levels of development.

[6] See Bazzi and Clemens (2013) for a discussion of the weak instrument diagnostics and consistent inferences.

The VAR methodology then allows examining simultaneously the effects in both directions. The authors find that shocks to the growth rate of income inequality have a negative effect on growth. Interestingly, they note that the relationship between inequality and growth varies over time and is sensitive to particular episodes in history, possibly because of common shocks to all states.

To the best of our knowledge, there are no attempts in the literature studying the effects of inequality on growth to look at the heterogeneity across countries. However, in his analysis Barro (2000) splits the sample into poorer and richer countries finding contradicting results, suggesting a potential bias caused by imposing homogeneity in slopes in the regression including both subsamples.[7] Similarly, De Dominicis et al. (2006) suggest restricting the analysis to a more homogeneous set of countries or to regions within a country. Reducing the sample to a more homogeneous one, or allowing for different slopes also addresses the omitted variable bias that arises from effects that vary across countries or groups of countries.

3 Data and Stylized Facts

Despite its acknowledged shortcomings, the Gini index remains the most commonly used measure of income inequality, mostly because of its wider data availability compared to other measures.[8] As explained in Solt (2016), there is a trade-off between comparability and coverage when it comes to the use of the Gini data. On the one hand, privileging comparability would mean using a low-coverage harmonized data from a single source or using data from a single basis of calculation (e.g., household-adult-equivalent disposable income). However, this option entails neglecting a lot of available information. On the other hand, favoring coverage would mean considering more of the available data and making adjustments to account for different basis of calculation.

The literature suggests different approaches or adjustments to exploit the available information, nevertheless they remain unsatisfactory. Halter et al. (2014), for example, simply ignore comparability issues. Following the recommendations of Deininger and Squire (1996), Milanovic (2013) includes dummy variables to adjust for different welfare definitions and equivalence scales. This fixed adjustment, however, is imperfect as it disregards that the relationship between income and income inequality can vary over time and across countries and could result in under- or over-estimation in specific country-years. Thus, these methods seem inferior to the one adopted by Solt (2016) in the construction of the Standardized World Income Inequality Database (SWIID).

The SWIID starts from the Gini indexes of the Luxemburg Income Study (LIS)—considered fully comparable data—and the Gini indexes in the source data. In the SWIID 5.1, the source data includes ten thousand Gini indexes from 1960 to 2014 based on all or nearly all the countries' population and for which there is sufficient information to identify the equivalence scale and welfare definition used in the calculation. Then, model-based multiple imputation is employed to estimate the missing observation of the LIS starting from the source data.[9] For each variable in the

[7] Barro (2000) also warns that poorer countries are likely to have a large measurement error in national accounts.

[8] The Gini index is often criticized because it is argued that it is more sensitive to the income of the middle classes than that of the extremes. Also, economies with similar incomes and Gini coefficients can have very different income distributions reflecting the fact that the Lorenz curve can have different shapes and still yield the same Gini index.

[9] The procedure consists of creating several categories of welfare definition and equivalence scale, constructing time- and country-varying ratios of indexes available in different categories, and generating the missing ratios using the least uncertain predictions of a set of regression models. Ratios for each category are finally combined into a final estimate by taking into account the estimate with the smallest uncertainty, or the mean of some estimates if the associated standard error is even smaller. If dramatic variations from one year to the next are detected, the proce-

SWIID, there are 100 multiply-imputed values to allow users to control for measurement error in the statistics. Version 5.1 of the SWIID represents an improved and enlarged version of the one used by Ostry et al. (2014), and addresses some of the issues raised in Jenkins (2015). Moreover, as shown in Solt (2016), the SWIID is successful in predicting the LIS data before their release, conferring some confidence in its cross-section and over-time comparability. Nonetheless, issues remain. Namely, the fact that correlations are mostly estimated over advanced economies and later applied to other countries, implicitly assuming that the same relationship holds across countries, and the interpolation over time.

Our dataset draws from the SWIID 5.1 for net Gini (where net refers to its measurement after taxes and transfers) and from the Penn World Tables (PWT) 9.0 for the real income variable. Net Gini (and not market Gini) is the relevant variable for the analysis as only post-redistribution income is what matters for individuals. [10] We use data on population from PWT 9.0 to compute real per capita GDP. Owing to the requirements of our estimation methodology, we drop some observations by applying the following criteria. First, we drop all non-continuous time series, remaining with the longest run for every country for which data is available. Second, if the remaining run is composed by less than 20 observations, we drop all observations for that country.[11] Third, we drop countries for which the net Gini is stationary to work with variables integrated of the same order.[12] After this procedure, we are left with a dataset of 77 countries spanning at least 20 years, for a total of 1,597 data points.

Figure 1 presents some key stylized facts. The top left panel shows the median net Gini along with the interquartile range over the period 1990-2010. The use of such period ensures that at least 95 percent of the countries in the sample are represented. The median income inequality slowly increased from 33 percent in 1990 to about 37 percent in the wake of the global financial crisis, after which it started to mildly decline to about 35 percent in 2010. The 25^{th} and 75^{th} percentiles present a somewhat similar trend. However, they reveal a wide variation across countries, with one fourth of them having a net Gini up to 25-30 percent and another fourth of the countries with a net Gini higher than 42-47 percent over the selected period.

We now turn to analyze such heterogeneity across regions and income groups. The top right panel shows that median net Gini is highest in Africa and the Western Hemisphere, and that it has been falling at a faster pace in the former. The more equal income distribution is observed in Europe, albeit the net Gini presents an upward trend. Asia and the Pacific and the Middle East and Central Asia experienced a similar rate of growth, but started from a higher level. In the case of the Middle East and Central Asia, such upward trend has been somewhat reversed since 2004. The middle left panel suggests that the median income inequality remained relatively lower but steadily increasing in advanced economies, while it was more volatile and fluctuating at compara-

dure allows observations to be informed by the estimates for surrounding years with a five-year weighted moving average applied to a variable generated through Monte Carlo simulations. Exceptions to the last step apply if dramatic changes are detected in the LIS series or are related to countries of Eastern Europe and former Soviet Union during the collapse of communist rule. The last step generates estimates after 1975 by interpolation.

[10] The net Gini, as opposed to the market Gini, excludes redistributive interventions such as expenditure programs, notably on education and health, transfers to the poorer segments of the population, and non-proportional taxes, among others. However, net Gini still includes a wide range of non-redistributive policies (e.g., job-training programs and capital-accounts regulations), that can shape the income distribution (Iversen and Stephens, 2008; Morgan and Kelly, 2013).

[11] Pedroni (2013) presents the remarkable small sample properties of the heterogeneous panel SVAR. In particular, it is shown that even in an extreme case with $N = 50$ and $T = 12$, where time series estimates become unreliable and possibly divergent, there is another non-divergent set of time series that contain a signal about the true dynamic structure. By correlating this signal cross-sectionally with a static observable and using the fitted estimates, the true structural dynamics of the individual panel are traced out despite the very short time series.

[12] The application of this criterion implies the loss of 9 countries, or 11 percent of the dataset obtained after applying the second criterion.

ble levels for emerging markets and low-income developing countries.

In the remaining panels, we take a look at the relationship between net Gini and real per capita GDP. The middle left panel shows the medians for the two variables. The two lines appear negatively correlated, and this is confirmed by the fractional polynomial regression and its confidence interval in the bottom left panel. However, this panel also reveals a large potential for heterogeneous dynamics in our sample. In fact, the same fractional polynomial regression for eight (or about 10 percent) randomly selected countries in the sample suggests that the strength and the direction of relationship might be different across countries.

Cross-section dependence caused by common shocks (e.g. the global financial crisis) may also be a concern. To display its potential in our dataset, for every country we select the year in which it experienced the highest net Gini growth rate, this time over the entire period covered by the data. Then, we plot in the bottom right panel the number of countries experiencing the highest net Gini growth rate for every year. Although this is a very preliminary evidence, the bars suggest that a bunch of countries experienced their highest net Gini growth rate in the same years, possibly reflecting common shocks.

4 Econometric Analysis

4.1 Methodology and Identification Strategy

As noted in Section 2, the relationship between income inequality and growth has proven difficult to analyze for a number of reasons. First, the two variables may be intertwined, such that income inequality has an impact on growth at the same time that growth has an impact on income inequality. The failure to account for this source of endogeneity may bias the results. Endogeneity may also arise because of omitted variables that are correlated either with income inequality or growth. Second, dynamics are likely to be present, in the sense that the feedback between income inequality and growth occurs gradually over time, and with different intensities over different time horizons. Third, there is no reason for which these dynamic relationships would be the same across countries, regions, or income levels. In fact, they may differ because of the economic structure, institutions, policies' inclusiveness, among others. Fourth, shocks that are common to all countries might induce cross-section dependence and omitted variable bias.

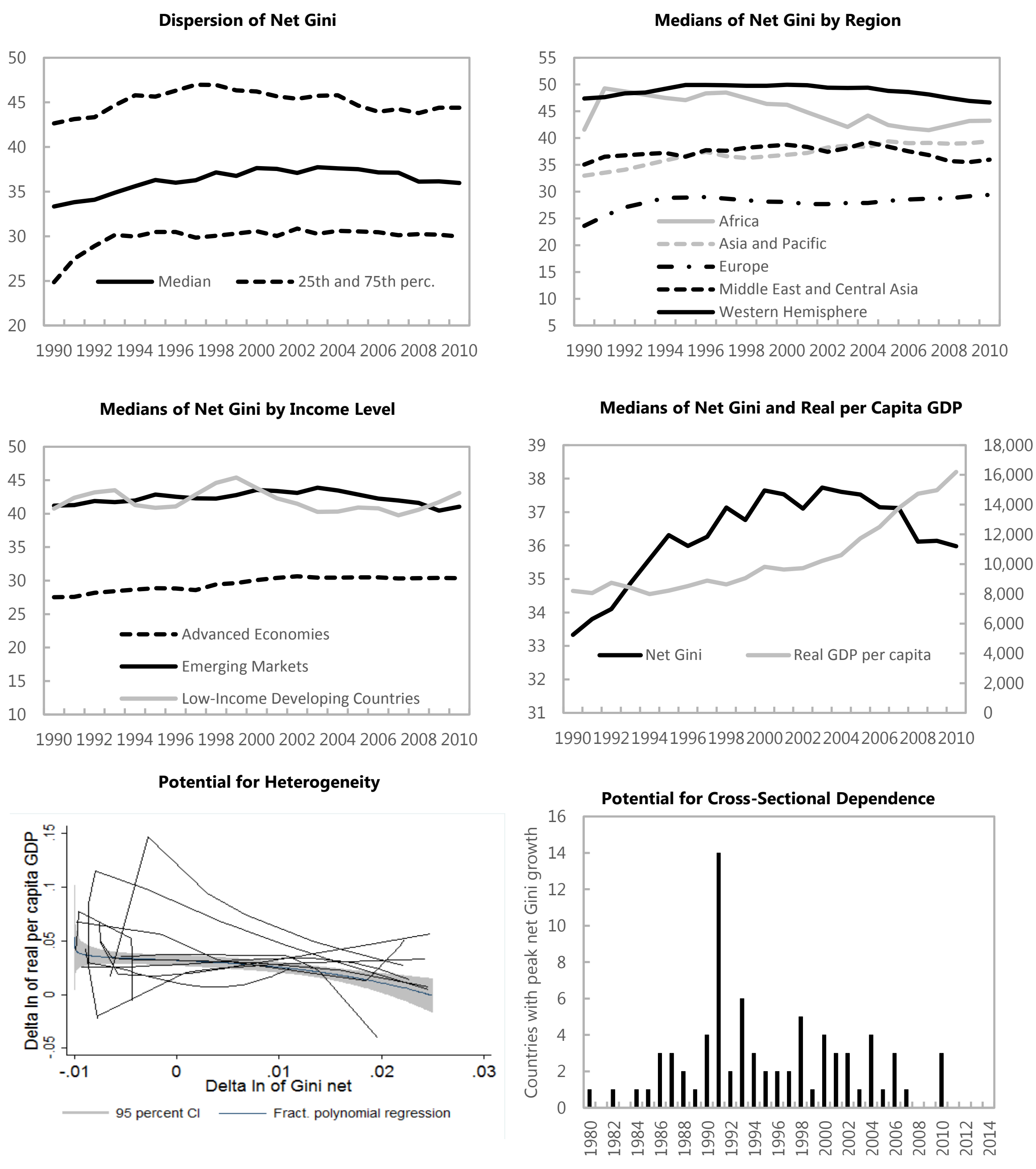
To deal with these complexities, we employ the heterogeneous panel SVAR model of Pedroni (2013).[13] Let $y_{i,t}$ be a vector of n endogenous variables (i.e., the log differences of real per capita GDP and net Gini) with country-specific time dimension $t = [1, ..., T_i]'$ for each member $i = [1, ..., M]'$ of an unbalanced panel. As a way to deal with country-fixed effects, we demean the data as $y^*_{i,t} = y_{i,t} - \bar{y}_i$, where $\bar{y}_i = T_i^{-1} \sum_{t=1}^{T_i} y_{i,t} \forall i$. Thus, we estimate the following equation:

$$B_i \Delta y^*_{i,t} = A_i(L) \Delta y^*_{i,t-1} + e_{i,t} \tag{1}$$

where $A_i(L)$ is a polynomial of lagged coefficients ($A_i(L) = \sum_{j=0}^{J_i} A^i_j L^j$) with country-specific lag-lengths J_i, A^i_j is a nxn matrix of coefficients, $e_{i,t}$ is a vector of residuals, and B_i is a nxn matrix of contemporaneous coefficients. J_i are selected based on the Schwartz Information Criterion to assure that residuals approximate white noise. As a way to control for the uncertainty associated to the net Gini measurement, we use the inverse of the standard deviation of the 100 multiply-imputed values as weights, thereby estimating by weighted least squares (WLS).

[13] The presentation of the technique draws from Pedroni (2013) and Góes (2016).

Figure 1: Stylized Facts



Notes: Top and middle panels are shown only over the period 1990-2010 to have at least 95 percent of the countries in the sample represented. Fractional polynomial regressions are shown for the sample along with the confidence interval and for eight randomly selected countries.
Source: Authors' calculations.

As noted, the dynamic relationship between income inequality and growth can differ across countries. Failing to account for such heterogeneity by simply pooling data as in conventional dynamic panel methods may result in inconsistent estimation and inference of the relationships.[14] Thus, the methodology employed in this paper relaxes the assumption of homogeneous dynamics among the members of the panel by estimating and identifying a reduced-form VAR for each country i:

$$\begin{aligned} B_1 \Delta y^*_{1,t} &= A_1(L)\Delta y^*_{1,t-1} + e_{1,t} \\ &\vdots \\ B_M \Delta y^*_{M,t} &= A_M(L)\Delta y^*_{M,t-1} + e_{M,t} \end{aligned} \tag{2}$$

from which we can obtain the idiosyncratic dynamics. To the extent that correlated omitted variables have a heterogeneous impact across countries, treating each country separately can also help in dealing with the bias generated by the omissions. Another advantage of relying on the heterogeneous approach is that between-countries comparability is less of an issue as each member of the panel is treated independently.

Then, we estimate an additional VAR using cross-section averages for each period, which would serve to capture to retrieve the common dynamics:

$$\bar{B}\Delta \bar{y}^*_t = \bar{A}(L)\Delta \bar{y}^*_{t-1} + \bar{e}_t \tag{3}$$

where $\bar{y}^*_t \equiv M^{-1}\sum_{i=1}^{M} y^*_{i,t}$.

We can now recover the SVAR representation, in which the reduced-form residuals in equation (2) and (3) are mapped into $u_{i,t} = B_i^{-1}e_{i,t}$ and $\bar{u}_t = \bar{B}^{-1}\bar{e}_t$, respectively. Thus, we calculate nM correlation coefficients to construct M diagonal matrices:

$$\Lambda_i = \begin{bmatrix} \rho(u^1_{i,t}, \bar{u}^1_t) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \rho(u^n_{i,t}, \bar{u}^n_t) \end{bmatrix} \tag{4}$$

where $\rho(u^n_{i,t}, \bar{u}^n_t)$ denote the correlation coefficients between structural residuals of the n^{th} endogenous variable for each country i. We are now able to decompose the composite shocks $u_{i,t}$ as:

$$u_{i,t} = \Lambda_i \bar{u}_{i,t} + \tilde{u}_{i,t} \tag{5}$$

where $\bar{u}_{i,t}$ are the common shocks, $\tilde{u}_{i,t}$ are the idiosyncratic shocks, and Λ_i are nxn diagonal matrices containing country specific loadings which account for the relative importance of common shocks. We then recover the matrices of composite responses to structural shocks for each country $R_i(L)$, and use the loading matrices estimated in equation (5) to disentangle the composite responses into responses to common shocks and responses to idiosyncratic shocks:

$$R_i(L) = \Lambda_i R_i(L) + (I - \Lambda_i \Lambda_i')R_i(L) \tag{6}$$

It should be noted that $R_i(L) = \bar{R}_i(L) + \tilde{R}_i(L)$, where $\bar{R}_i(L) \equiv \Lambda_i R_i(L)$ and $\tilde{R}_i(L) \equiv (I - \Lambda_i \Lambda_i')R_i(L)$. Thus, we rely on the cross-section distribution of $R_i(L)$, $\bar{R}_i(L)$, and $\tilde{R}_i(L)$ to present selected descriptive statistics of the impulse response functions (IRF), such as their medians, averages, and interquartile ranges. Finally, we also calculate standard errors of medians through a re-sampling simulation with 500 repetitions.

The structural representation of the VAR requires an identification strategy. Rather than imposing economic relationships, we rely on the Cholesky decomposition. The ordering of the variables

[14] See Pesaran and Smith (1995).

Table 1: Panel Unit Root Tests

	Levels		First differences	
	Intercept	Intercept and trend	Intercept	Intercept and trend
Levin, Lin, and Chu (2002)				
Ln net GINI	-0.617	-1.110	-12.488**	-6.933**
Ln per capita real GDP	-1.419	-2.58	-23.297**	-19.438**
Im, Pesaran, and Shin (2003)				
Ln net GINI	-0.234	-2.200	-18.833**	-15.324**
Ln per capita real GDP	8.137	-1.070	-22.258**	-19.927**
Maddala and Wu (1999)				
Ln net GINI	-0.023	-0.067	-16.687**	-14.044**
Ln per capita real GDP	7.685	-0.524	-20.329**	-17.768**

Source: Authors' calculations.
Notes: The null hypothesis of the panel unit root tests is that all panels contain a unit root. Fixed effects are always included. The Schwartz Information Criterion is used to select the optimal lag length. ** and * next to a number indicate statistical significance at 1 and 5 percent, respectively.

is based on the following argument. Since income is an input in the calculation of income inequality, any change in income will have a contemporaneous effect on income inequality. At the same time, changes in income inequality are likely to have only a delayed impact on income. Therefore, the identification strategy rests on lag restrictions rather than economic ones, implying that changes in the net Gini do not have contemporaneous (i.e., within a year) effect on growth.

4.2 Integration, Cointegration, and Convergence

As a first step, we analyze the stationarity properties of the log of Gini net and the log of per capita real GDP by employing the tests by Levin et al. (2002) (LLC thereafter), Im et al. (2003) (IPS thereafter), and Maddala and Wu (1999) (MW thereafter).[15] These tests differ in that LLC treats the parameter of interest as common across countries and focuses on the within-country dimension, IPS treats the parameter of interest as varying across countries and focuses on the between-country dimension, and MW treats all the parameters as potentially varying across countries and tests by pooling significance values across members of the panels.

We perform the tests on the (log) levels and the (log) first differences of the series, both including the intercept only and the intercept and the trend. As expected, the results of the different tests reported in Table 1 suggest that levels are non-stationary as the null hypothesis of unit root cannot be rejected, irrespective of the inclusion of a time trend. After first differencing, the series result stationary.

The presence of unit roots in levels warrants the investigation of the existence of a long-run relationship between (the log of) Gini net and (the log of) per capita real GDP. We employ the tests by Kao (1999) and Pedroni (1999, 2004). The former treats all parameters except fixed effects as homogeneous, and therefore is mispecified if heterogeneous dynamics exist. The latter treat all pa-

[15] One could argue that since the net Gini is bounded between zero and one, it cannot contain a unit root. While this is true for long time series, there is still the possibility that the series is locally non-stationary. One way to address this possible critique would be to apply the logistic transformation $y_{i,t} = ln[n_{i,t}/(1 - n_{i,t})]$, where $n_{i,t}$ is the net Gini. Applying such transformation, however, does not affect any of the results presented in this paper.

Table 2: Panel Cointegration Tests

	Intercept	Intercept and trend
Kao (1999)		
ADF	-3.109**	
Pedroni (1999, 2004)		
Panel *v*-Statistic	0.833	3.921**
Panel *ρ*-Statistic	-1.681*	0.710
Panel *t*-Statistic (non-parametric)	-6.224**	-2.252*
Panel *t*-Statistic (parametric)	-3.774**	-5.022**
Group *ρ*-Statistic	1.168	4.185
Group *t*-Statistic (non-parametric)	-4.191**	0.674
Group *t*-Statistic (parametric)	-5.782**	-5.662**

Source: Authors' calculations.
Notes: The null hypothesis of the panel cointegration tests is that there is no cointegration between ln Gini net and ln real per capita GDP. For Pedroni (1999, 2004), we present the versions of the statistics that are not weighted by the member specific long-run conditional variances. The Schwartz Information Criterion is used to select the optimal lag length. ** and * next to a number indicate statistical significance at 1 and 5 percent, respectively.

rameters as heterogeneous across countries and include within- and between-country dimension tests, both parametric and non-parametric.

The results of the cointegration tests reported in Table 2 suggest the existence of a long-run relationship between inequality and income per capita. In particular, in the majority of cases and irrespective of the inclusion of a trend, both homogeneous and heterogeneous tests reject the null hypothesis of no cointegration, providing evidence of comovements between the series in the long run.

We now turn to investigate if income inequality data show long-run convergence or divergence across countries. Previous studies of income inequality convergence include Sala-i-Martin (2006) and Chotikapanich et al. (2012), that use non-parametric or parametric methods to estimate world inequality indexes and assess convergence by looking at the trend in the world distribution of income. We take a different approach. Borrowing from the literature on income convergence (Evans, 1998; Pedroni and Yao, 2006), we define convergence as the reduction of the inequality gap between any pair of countries. As noted by Pedroni and Yao (ibid.), an empirical formalization differs from the graphical inspection in that it requires that properties of the data must be consistent with the fact that differences are eliminated eventually. This is something harder to detect using graphical analysis, where we can only observe if differences become smaller at any point in time.

Empirically testing for convergence is equal to testing for the stationarity of the difference between the levels of the series for any pair of countries.[16] This would imply that the series are cointegrated. Formally:

$$(y_{i,t} - y_{j,t}) \sim I(0) \forall (i,j) \tag{7}$$

where y is (the log of) net Gini, and (i, j) are country pairs. Similarly, testing for convergence across multiple countries requires that every possible country pair within the sample exhibits con-

[16] If stationary differences between countries have nonzero mean, convergence is said conditional, as it depends on the country-specific fixed effects.

Table 3: Convergence in Income Inequality

	Countries	Net Gini: Im, Pesaran, and Shin (2003)	Net Gini: Maddala and Wu (1999)	Market Gini: Im, Pesaran, and Shin (2003)	Market Gini: Maddala and Wu (1999)
Full sample					
$(y_{it}-\overline{y}_t)\ \forall i$	77	-3.939**	-3.392**	-2.506*	-3.204*
Regions					
$(y_{it}-\overline{y}_t)\ \forall i \in AFR$	11	-0.461	-0.430	-0.460	-0.473
$(y_{it}-\overline{y}_t)\ \forall i \in APD$	12	-0.905	-0.903	-1.649*	-1.624
$(y_{it}-\overline{y}_t)\ \forall i \in EUR$	29	-4.513**	-3.706**	-3.686**	-3.275**
$(y_{it}-\overline{y}_t)\ \forall i \in MCD$	9	-4.296**	-4.086**	-4.043**	-3.582**
$(y_{it}-\overline{y}_t)\ \forall i \in WHD$	16	-1.184	-0.872	-0.432	-0.091
Income groups					
$(y_{it}-\overline{y}_t)\ \forall i \in AE$	28	-3.705**	-2.665**	-4.864**	-4.470**
$(y_{it}-\overline{y}_t)\ \forall i \in EM$	36	-0.890	-0.595	-2.404	-1.927
$(y_{it}-\overline{y}_t)\ \forall i \in LIDC$	13	-0.463	-0.523	-1.260	-1.254

Source: Authors' calculations.
Notes: The null hypothesis of the unit root tests is that all panels contain a unit root, i.e. there is no convergence. Fixed effects are always included. The Schwartz Information Criterion is used to select the optimal lag length. ** and * next to a number indicate statistical significance at 1 and 5 percent, respectively.

vergence. However, Evans showed that if the the condition in equation 7 holds, then:

$$N^{-1}\sum_{j=1}^{N}(y_{i,t}-y_{j,t}) \sim I(0) \tag{8}$$

Thus:

$$N^{-1}\sum_{j=1}^{N}(y_{i,t}-y_{j,t}) = N^{-1}\sum_{j=1}^{N}y_{i,t} - N^{-1}\sum_{j=1}^{N}y_{j,t} = y_{i,t}-\bar{y}_t \tag{9}$$

So:

$$(y_{i,t}-y_{j,t}) \sim I(0)\forall(i,j) \iff (y_{i,t}-\bar{y}_t) \sim I(0)\forall i \tag{10}$$

In other words, convergence requires that $(y_{it}-\bar{y}_t)$ is stationary for each country in the sample. Therefore, the null hypothesis of nonconvergence can be resumed in the unit root null of panel unit root tests. As in Pedroni and Yao (2006), we employ the IPS and the MW panel unit root tests, leveraging on the fact that the autoregressive parameter in the income inquality differentials is allowed to vary across countries.

Table 3 presents the results of the tests of conditional convergence for (the log of) net Gini (and market Gini). Beyond the full sample, we also consider regions and income groups to investigate the existence of convergence clubs. The first column reports the Augmented Dickey-Fuller t-statistic of the IPS test along with its significance level. For the full sample, we reject the null hypothesis of unit root presence, suggesting convergence. When looking at the t-statistic for every country, the majority of them cannot reject the null hypothesis at 5 percent level. However, the signal from the latter is not neutral as it lies on the tail of the distribution, and when combined, is sufficient to warrant the rejection of the null hypothesis for the panel as a whole. This finding is supported by the Fisher statistic of the MW test.

We then analyze convergence across regions and income groups. The fact that countries are generally converging to the average for the sample does not exclude the possibility that countries are also converging to, for example, a regional average. The results of the unit root tests suggest that convergence is occurring in Europe and in the Middle East and Central Asia, as well as in advanced economies. Such findings are once again corroborated by the results of the MW test. While these conclusions are in general valid also in the case of market Gini, the null hypothesis for the full sample is rejected at a higher significance level both for IPS and MW, suggesting that the signal is weaker when redistribution effects are not netted out.

4.3 (Heterogeneous) Dynamics

We now present the IRF for the full sample. Figure 2 reports the median, the average, as well as the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile responses of real per capita GDP growth and net Gini growth among the 77 countries in the sample over a 12-year horizon. As noted by Góes (2016), this is an informative way of presenting the results, which would not be available if we had imposed homogeneity in the estimation. Particularly, we would not be able to know how many countries were actually behaving as the average dynamics, and inference would be biased in presence of heterogeneity (Pesaran and Smith, 1995).

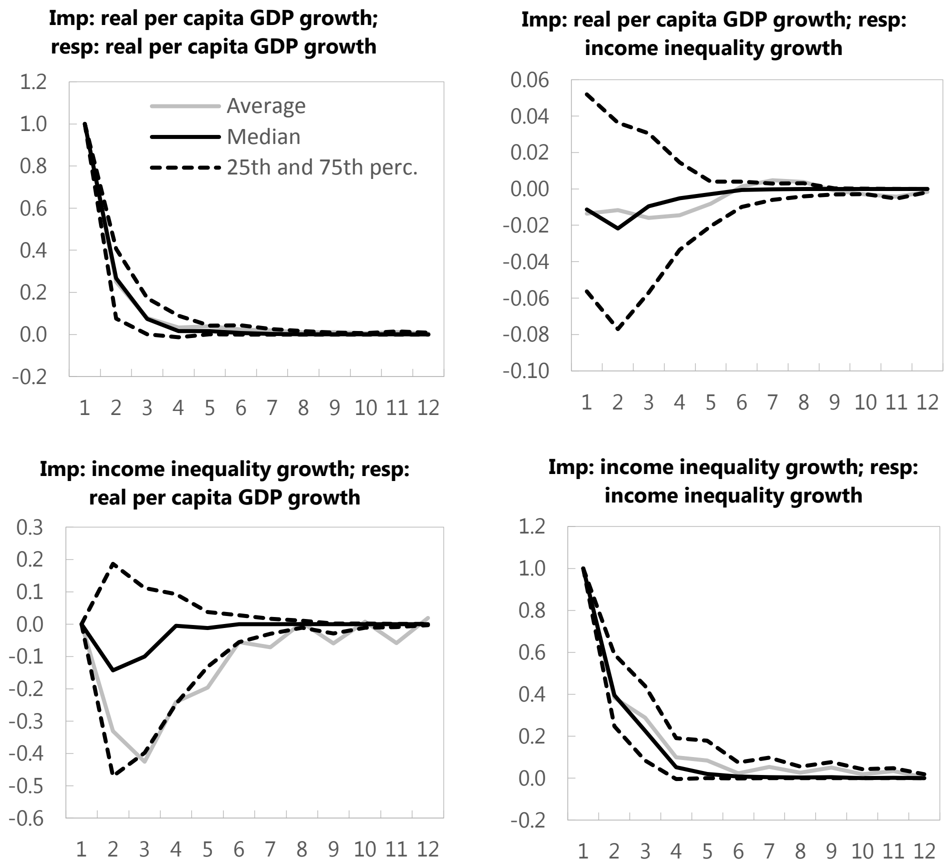
We find a large heterogeneity in the results. While the median impact of income inequality growth on real per capita GDP growth is negative and lasts for about two years, this is by no means true for all countries, as shown in the bottom left panel. Indeed, the results for the 25$^{\text{th}}$ percentile show that there is a subset of countries for which the estimated impact is actually positive and longer lasting. Conversely, the results for the 75$^{\text{th}}$ percentile suggest that the negative impact is about four times larger than the negative median response, and takes longer to die out. Importantly, the average response can be quite noisy in presence of heterogeneous slopes and dynamics. For the median response, a one percentage point (pp) shock in the growth rate of income inequality is associated with a 0.14 pp fall in real per capita GDP growth.

The results in Figure 2 are also informative regarding the impact of real per capita GDP growth on income inequality growth. The top right panel shows that the median response is negative and reverts to about zero after two years. However, once again, the variation across countries is considerable, as shown by the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile. The effects of the shocks in real per capita GDP growth and income inequality growth on themselves are positive as expected, and much more homogeneous (top left and bottom right panels).

As an illustration of the large heterogeneity, we present the cases of two different countries at the tails of the distribution. Figure 3 shows the estimated response of real per capita GDP to a shock in income inequality over a 12-year horizon for Nigeria and Finland. For Nigeria, the response is negative and large. It reaches its maximum magnitude three years after the shock (eight pp fall in real per capita GDP growth in response to a pp change in the net Gini) and still shows negative values 11 years after the shock. For Finland, the response is positive over the entire horizon considered, reaching its maximum magnitude one year after the shock (1.2 pp). While single-country results might be noisy, they still provide a suggestive picture of the heterogeneity in the strength of the effect.

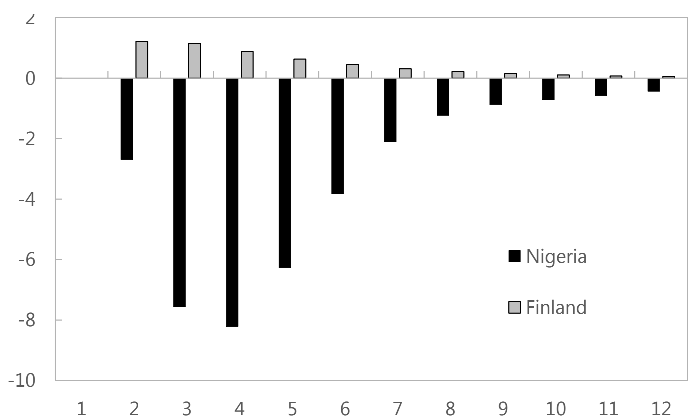
Our results so far suggest that the sign and the strength of the relationship between the two variables varies widely across countries. We now turn to investigate if such heterogeneity persists across regions (Africa, Asia and Pacific, Europe, Middle East and Central Asia, and Western Hemisphere) and income levels (advanced economies, emerging markets, and low-income developing

Figure 2: Composite IRFs



Source: Authors' calculations.

Figure 3: Composite Response of Real per Capita GDP Growth to Income Inequality Growth for Nigeria and Finland



Source: Authors' calculations.

countries).[17]

Figure 4 compares the median, the average, as well as the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile response of real per capita GDP growth to a shock in income inequality growth. With the exception of Asia and Pacific (which shows an erratic behavior) and the advanced economies (which present a mildly positive reaction in the second year), the median response remains negative. In terms of magnitude, the negative response to a one pp shock in the net Gini is largest in the Middle East and Central Asia and the Western Hemisphere (about -0.3 pp during the second and the third year after the shock, respectively), and in emerging economies (-0.4 pp one year after the shock). Even at the country group level, the median response can be deceiving.

While looking at the median gives an idea of the typical country in the subset, the interquartile range is telling about the dispersion. As in the case of the full sample, the dispersion is vast across regions and income groups. Only in Western Hemisphere the 25$^{\text{th}}$ percentile answer is negative, in all other cases it remains positive, suggesting that heterogeneity is present even at regional level. However, the extent of the positive effect is as large as the full sample or larger only for Asia and Pacific, the Middle East and Central Asia, and advanced economies; in all other cases, the magnitude is reduced.

We now move to look at the statistical significance of the median response of real per capita GDP growth to shocks in income inequality growth, both for the full sample and for sample subsets. Figure 5 reports the median response along with the associated 95 percent bootstrapped confidence intervals. The median effect is significant for the full sample at five percent level. Converted in index numbers and real PPP terms, these responses suggest that for the median country in the sample, an increase in the net Gini by 0.3 decreases per capita income by $24. However, a look at the regional and country group dynamics suggests that the effect remains significant only in the Middle East and Central Asia, the Western Hemisphere, and in emerging markets. In contrast, we cannot find any significant and positive effect in any sample subsets.

Interestingly, also the time needed for the effect of the shock to phase out is heterogeneous across country groups. For the full sample, the median response dies out three years after the shock. However, in emerging economies, it takes longer, as the median response reaches zero only five years after the shock. In the Middle East and Central Asia and the Western Hemisphere, the median response is significant two years after the shock and settles in two to three years.[18]

Figure 6 reports the corresponding median response of the level of real per capita GDP to shocks in income inequality growth for the full sample and sample subsets. To ease visualization, the median response is presented in terms of percent deviation from the steady state.[19] For the full sample, real per capita GDP is about 0.3 percent lower with respect to the steady state after a one pp shock in income inequality growth. The effect is more sizable in the Middle East and Central Asia

[17] Countries are classified in regional groups according to the IMF regional departments and in income levels according to the April 2016 IMF World Economic Outlook. See Appendix A for further details on country groups.

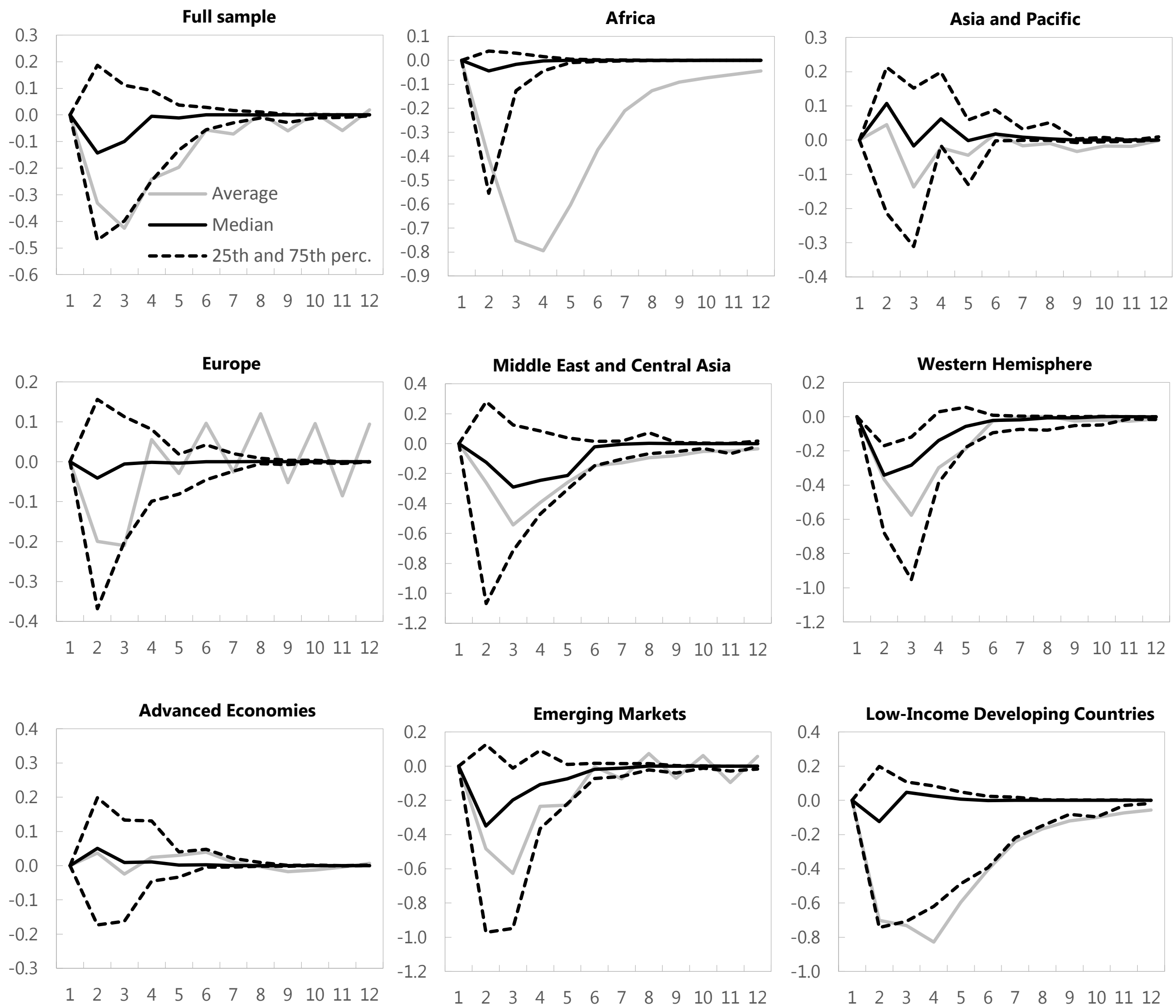
[18] Appendix B reports the median accumulated responses. Unsurprisingly, the long-run effect is significant only for the full sample (0.4 pp after a one pp shock in income inequality growth), and for Middle East and Central Asia (1.3 pp), the Western Hemisphere (1.4 pp), and emerging markets (0.9 pp).

[19] The percent deviation from the steady state ψ is calculated as:

$$\psi_{g,t+1} = \frac{y_{g,t}(1 + \dot{y}_g + IRF_{g,t+1}/100) - y^{ss}_{g,t}(1 + \dot{y}_g)}{y^{ss}_{g,t}(1 + \dot{y}_g)} 100$$

where the starting level (i.e., only at time t) of real per capita GDP y for the sample group g is equal to the median real per capita GDP for the sample group g, and is also equal to the steady state real per capita GDP y^{ss}; $\dot{y}_g$ is the median growth rate of real per capita GDP; and $IRF_{g,t+1}$ is the estimated IRF coefficient for the group g at time $t + 1$.

Figure 4: Composite Response of Real per Capita GDP Growth to Income Inequality Growth by Region and Income Level



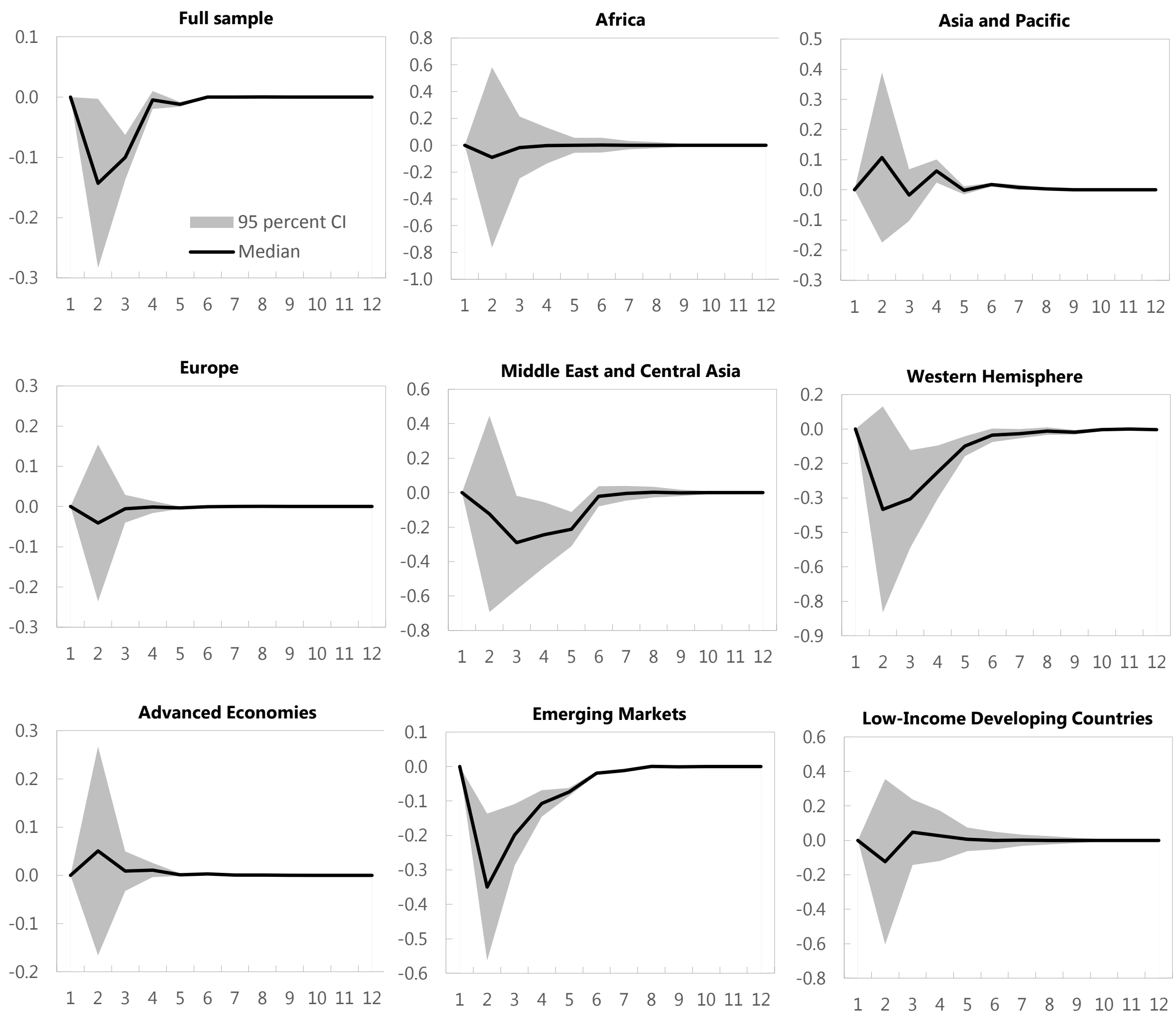
Source: Authors' calculations.

(0.9 percent), the Western Hemisphere (1 percent), and in emerging markets (0.8 percent).

Finally, we decompose the responses into those associated to idiosyncratic and common shocks.[20] As noted, loading factors contain information about the relative importance of common shocks for each country and can be used to decompose the aggregate response. Thus, in Figure 7 we present such decomposition for the response of real per capita GDP growth to income inequality growth. The results indicate that the great majority of the response is explained by idiosyncratic shocks.

[20] As noted by Yepes et al. (2015), another way to think about an idiosyncratic shock is to consider any shock which mostly affects only a single country, regardless of whether the geographic origin is specifically within or outside the country. Similarly, a shock which predominantly affects multiple countries, regardless of the specific geographic origin, is picked up as a common shock.

Figure 5: Composite Response of Real per Capita GDP Growth to Income Inequality Growth by Region and Income Level and Confidence Intervals

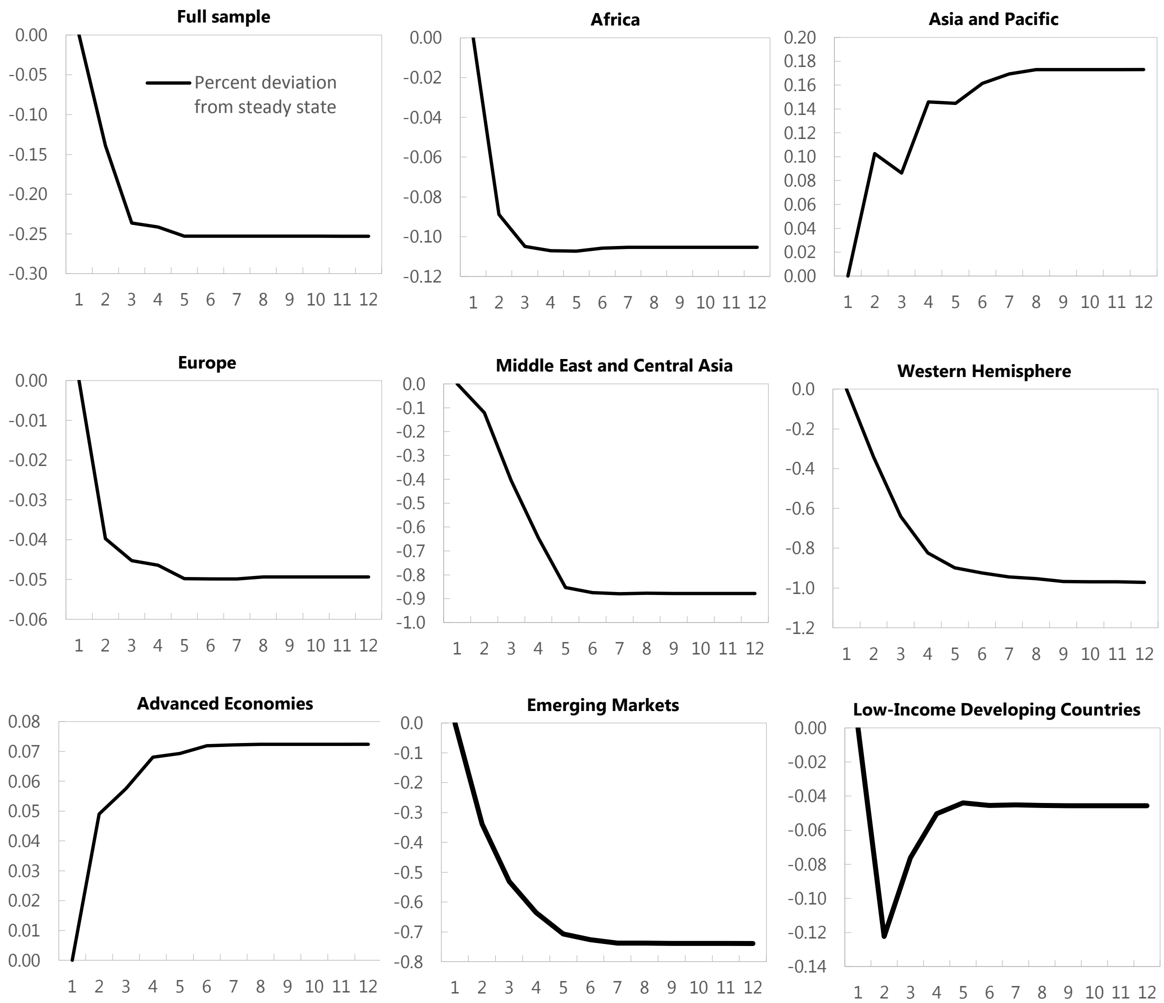


Notes: Confidence intervals are calculated from a resampling simulation with 500 repetitions.
Source: Authors' calculations.

4.4 A Look into the Relationship's Determinants

In order to explore the determinants of the variation in the IRF, we examine the cross-section association between certain country characteristics and the strength of the IRF. In particular, we test whether a better institutional framework could lead to a smaller negative effect of income inequality on growth. The benefits of a bad institutional environment tend to accrue to the more powerful higher income groups through the subversion of legal, regulatory, and political institutions, making them richer and worsening inequality (Glaeser et al., 2003 and Batabyal and Chowdhury, 2015). Thus, a higher institutional quality could soften the negative impact of inequality on growth. Another channel could be the one for which governments characterized by high institutional quality are often the more meritocratic ones, which provide incentives to compete and

Figure 6: Composite Response of Real per Capita GDP to Income Inequality Growth by Region and Income Level



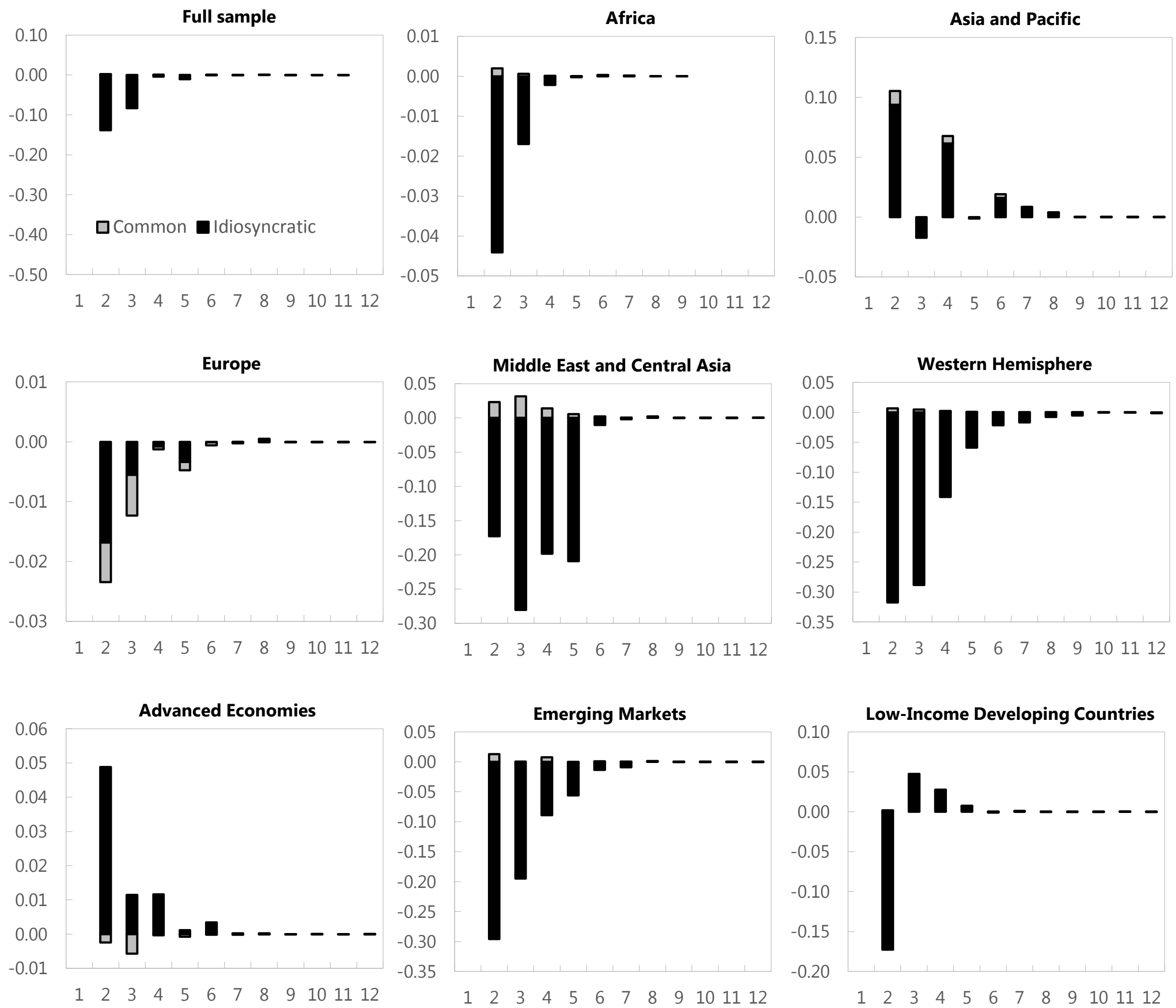
Source: Authors' calculations.

increase productivity, hence attenuating the deleterious effect of income inequality on economic activity. Also, we test if higher education has any effect on the extent to which income inequality reduces economic activity. One of the reasons this may happen is because better educated people may value education more strongly. Thus, even if the distribution of income becomes more unequal, they will refrain from reducing investment in education, which would contribute to mute the effect of income inequality on economic growth.

There are several indicators of institutional quality or development. We consider here a subset composed by control of corruption, rule of law, government effectiveness, and political stability from the Worldwide Governance Indicators and the Economic Freedom of the World Index from the Fraser Institute. As for education, we use years of schooling from the Barro-Lee database and education expenditure from the IMF's Fiscal Affairs Department database.

Measuring the impact of these variables on the strength of the relationship using the IRF coefficients is complicated by the fact that the response coefficients typically vary period by period,

Figure 7: Idiosyncratic and Common Median Response of Real per Capita GDP Growth to Income Inequality Growth by Regions and Income Levels



Notes: Median responses are decomposed into idiosyncratic and common shocks by using the loading factors.
Source: Authors' calculations.

implying that no single number provides an unambiguous measure of the size of the response. Accordingly, we examine the magnitude of each of the responses one to four years after the shock (which correspond to the years where most countries present a significant marginal response), as well as by the magnitude of the average response coefficient and the accumulated response coefficient over a 12-year horizon.[21] More formally, we estimate the following equation with an OLS estimator:

$$|IRF_{i,t}| = \alpha + \beta IQ_i + \gamma EDUC_i + \epsilon_i \tag{11}$$

where $|IRF_i^t|$ is the absolute value of, alternatively, the IRF coefficient at time t, the average response coefficient, or the accumulated response coefficient for country i; IQ is the institutional quality variable; $EDUC$ is the proxy for education attainment; and ϵ is a white noise error term.

[21] Mishra et al. (2014) use a similar approach.

As a first step, we examine the bivariate correlation between the IRF coefficients and the potential correlates. The scatter plots of Figure 8 only shows the relationship for the IRF coefficients one year after the shock and the accumulated ones for space reasons. The signs of almost all the bivariate correlations (13 out of 14) are consistent with the hypotheses outlined above. Higher institutional quality and more years of schooling (or higher education spending) are associated with a smaller reduction in economic growth in response to an income inequality shock. With the exception of the correlation between years of schooling and the accumulated IRF coefficients, the bivariate regression coefficient is always significant.

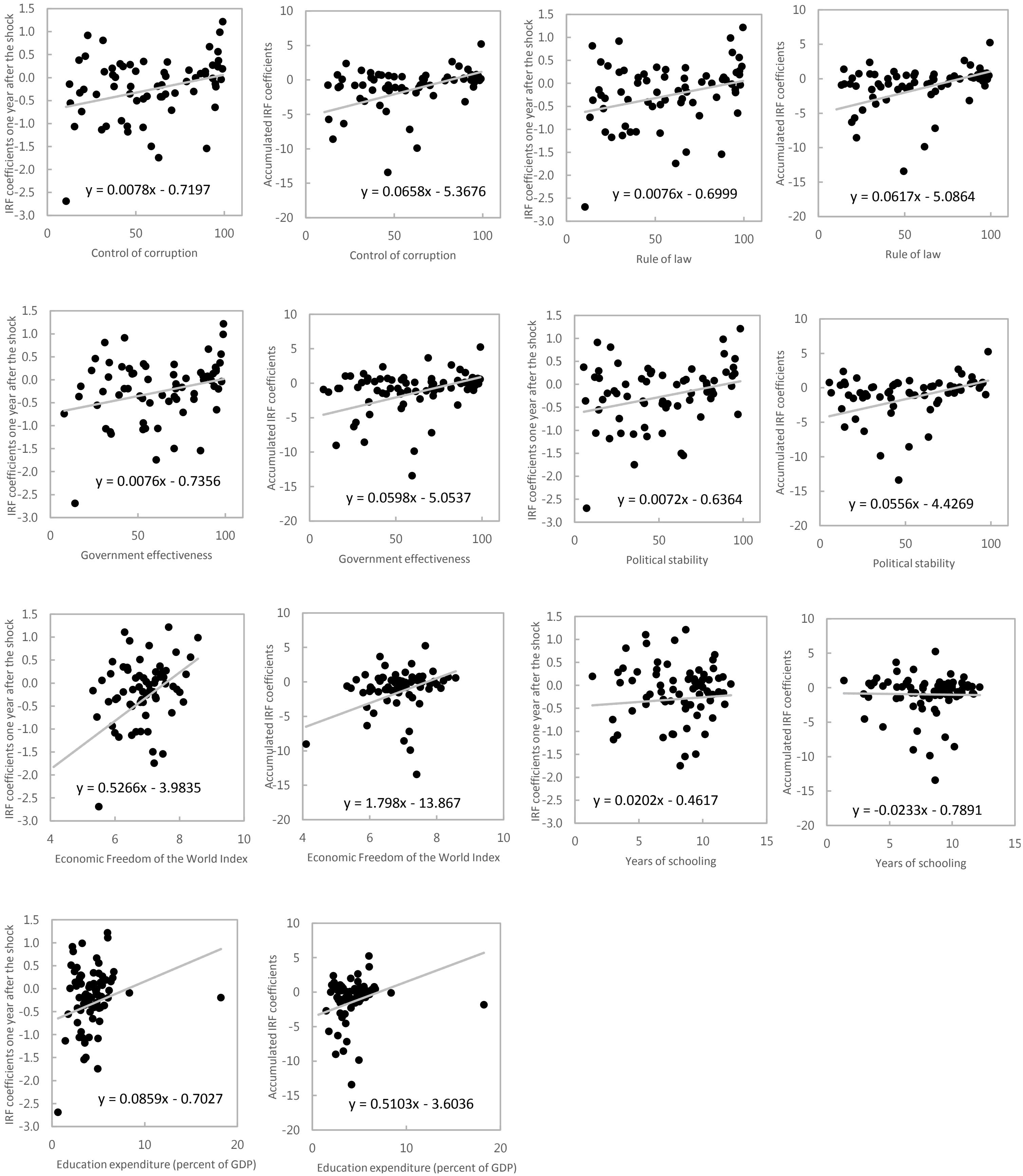
The multivariate regression results are presented in Table 4. Each panel reports the results of the regressions of the IRF coefficients one to four years after the shock, the average coefficients over the 12-year horizon, and the accumulated IRF coefficients on a specific institutional quality variable and years of schooling.[22] We report here the results using years of schooling as a proxy for education attainment. The results are generally robust to substituting years of schooling with education expenditure.

The results for control of corruption, rule of law, and government effectiveness are generally consistent with the bivariate correlations in Figure 8. In the case of control of corruption and rule of law, we find that there is a negative and significant relationship at nearly all the horizons considered, as well as for the average and accumulated IRF coefficients. In the case of government effectiveness, the result holds only for the average and accumulated IRF coefficients.[23] This is particularly relevant considered that we have only 75 observations in a regression that is designed to explain the cross-section values of very heterogeneous estimated parameters. At the same time, we find little or no evidence suggesting that stronger political stability, a higher value of the Economic Freedom of the World Index, and more years of schooling act as absorbers of income inequality shocks on economic activity.

[22] Institutional quality variables are highly correlated. To avoid multicollinearity, they are entered in the specification at a time.

[23] As a robustness test, we estimate the same specifications with a WLS estimator, using the inverse of the standard error of estimated IRF coefficients as weights. The results are similar to the ones presented in Table 4, suggesting that improvements in controlling corruption, rule of law, and government effectiveness reduce the impact of inequality on growth.

Figure 8: Bivariate Correlations



Source: Authors' calculations.

Notes: The minimum of the vertical axis for the accumulated IRF coefficients has been set to -20 to ease the visualization of the data. This leaves out Nigeria, which is however part of the sample to calculate the regression coefficient.

Table 4: Multivariate Correlations

	(1) IRF coefficients 1 year after the shock	(2) IRF coefficients 2 years after the shock	(3) IRF coefficients 3 years after the shock	(4) IRF coefficients 4 years after the shock	(5) Average IRF coefficients	(6) Accumulated IRF coefficients
Control of corruption	-0.005	-0.006	-0.004*	-0.003*	-0.002**	-0.028**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.001)	(0.012)
Years of schooling	-0.024	0.022	0.012	0.012	0.008	0.092
	(0.046)	(0.043)	(0.027)	(0.018)	(0.011)	(0.137)
Constant	1.070**	0.702**	0.486**	0.295**	0.224**	2.691**
	(0.332)	(0.310)	(0.190)	(0.131)	(0.082)	(0.979)
Observations	74	74	74	74	74	74
R-squared	0.040	0.032	0.043	0.045	0.072	0.072
Rule of law	-0.006	-0.006	-0.004*	-0.003*	-0.002**	-0.029**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.001)	(0.013)
Years of schooling	-0.016	0.024	0.013	0.014	0.009	0.108
	(0.047)	(0.044)	(0.027)	(0.019)	(0.012)	(0.140)
Constant	1.062**	0.683**	0.473**	0.286**	0.217**	2.606**
	(0.329)	(0.309)	(0.190)	(0.131)	(0.081)	(0.975)
Observations	74	74	74	74	74	74
R-squared	0.047	0.031	0.042	0.045	0.072	0.072
Government effectiveness	-0.004	-0.005	-0.004	-0.003	-0.002**	-0.030**
	(0.005)	(0.005)	(0.003)	(0.002)	(0.001)	(0.014)
Years of schooling	-0.025	0.019	0.012	0.013	0.010	0.118
	(0.050)	(0.047)	(0.029)	(0.020)	(0.012)	(0.149)
Constant	1.059**	0.695**	0.488**	0.299**	0.227**	2.720**
	(0.333)	(0.313)	(0.192)	(0.132)	(0.082)	(0.987)
Observations	74	74	74	74	74	74
R-squared	0.033	0.017	0.030	0.030	0.061	0.061
Political stability	-0.006	-0.005	-0.003	-0.002	-0.001	-0.016
	(0.004)	(0.004)	(0.002)	(0.002)	(0.001)	(0.013)
Years of schooling	-0.018	0.019	0.007	0.008	0.002	0.020
	(0.046)	(0.043)	(0.027)	(0.019)	(0.012)	(0.140)
Constant	1.040**	0.665**	0.461**	0.278**	0.206**	2.475**
	(0.328)	(0.309)	(0.190)	(0.131)	(0.083)	(0.995)
Observations	74	74	74	74	74	74
R-squared	0.047	0.026	0.030	0.024	0.026	0.026
Economic Freedom of the World Index	-0.321**	-0.107	-0.103	-0.053	-0.049	-0.587
	(0.159)	(0.154)	(0.094)	(0.065)	(0.041)	(0.487)
Years of schooling	-0.004	0.001	0.003	0.004	0.003	0.033
	(0.050)	(0.049)	(0.030)	(0.020)	(0.013)	(0.154)
Constant	2.886**	1.291	1.037*	0.560	0.467*	5.605*
	(0.940)	(0.914)	(0.557)	(0.382)	(0.240)	(2.882)
Observations	67	67	67	67	67	67
R-squared	0.086	0.010	0.024	0.012	0.026	0.026

Source: Authors' calculations.
Notes: Robust standard errors are reported. ** and * next to a number indicate statistical significance at 1 and 5 percent, respectively.

4.5 Robustness

The estimation methodology employed in this paper should address to the best extent possible issues related to endogeneity caused by reverse causality, cross-country comparability, bias induced by assuming homogeneity across countries, and cross-section dependence. Nonetheless, other issues common to the literature may persist. First, measurement error may bias the estimates. Second, since the adopted identification strategy is based on some assumptions, these may be driving the thrust of the results. Third, endogeneity may be present due to some time-varying omitted variable with an idiosyncratic dynamic and correlated to any of the endogenous variables in the empirical model. Fourth, dropping countries for which income inequality data is stationary may bias the results.

We discuss here the results from a battery of robustness tests which address some of these concerns. Figure 9 compares the marginal response of real per capita GDP growth to income inequality growth of the baseline model to the one obtained from some variations of it. As a first step, we substitute net Gini with market Gini. This is not strictly a robustness test as post-redistribution data measure the effects of various policy interventions, however they can still provide insights about the strength of the relationship.[24] Results in the top left panel confirm the negative and significant relationship between the variables.

While our identification strategy is based on the fact that income inequality is an input in the calculation of income inequality and therefore it is contemporaneously determined, we test the robustness of the results by inverting the ordering of the Cholesky decomposition. As shown in the top middle panel, such modification of the identification scheme allows changes in net Gini growth to have an immediate effect on real per capita GDP growth. Although with a different timing, such modification of the specification still replicates the negative and significant effect observed in the baseline results.

Similar to other papers employing different methodologies, we then try to add to our specification one variable at a time that is potentially correlated with any of the endogenous variables.[25] As a first test, we add redistribution from the SWIID 5.1 in a PSVAR with real per capita GDP growth and net Gini. The top right panel confirms the negative and significant response of income inequality growth on economic growth, even though with a somewhat delayed effect. The same experiment is repeated in the bottom left panel by adding real investment growth to the specification, without substantial changes to the estimates. Also, we add terms of trade growth and the results in the bottom center panel confirm that the marginal response is consistent with that estimated using the baseline specification.[26]

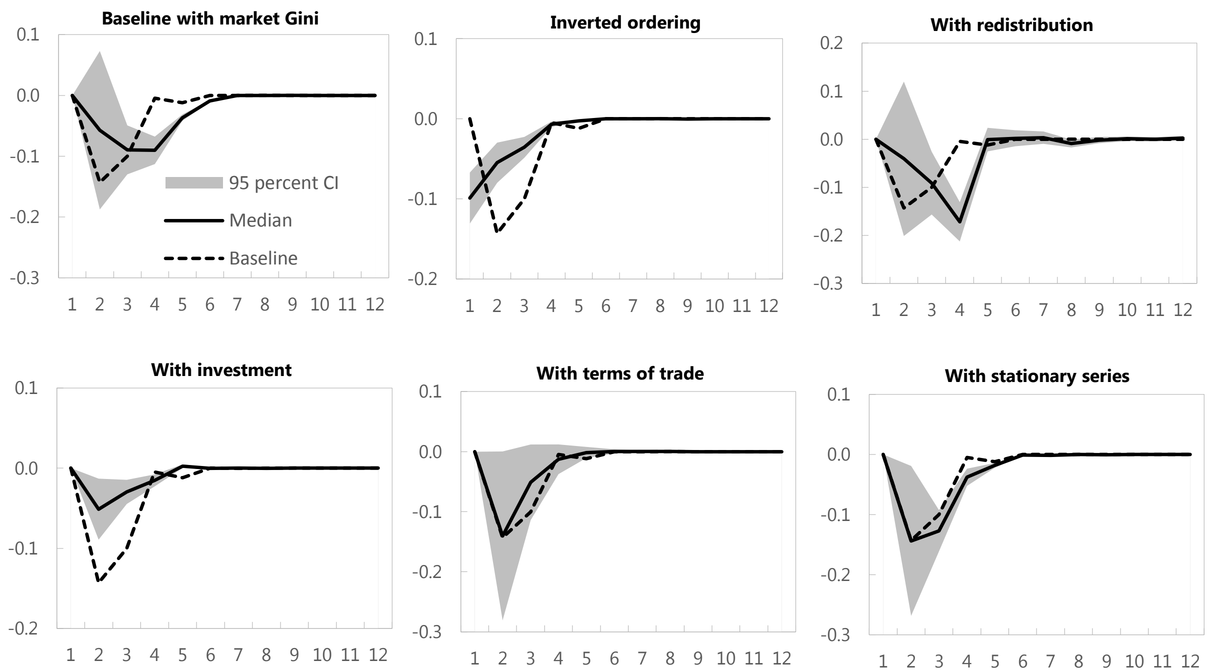
Finally, we augment the dataset by adding data for the 9 countries presenting stationary series of net Gini and rerun the PSVAR. The results for the augmented dataset comprising 86 countries are shown in the bottom right panel. The estimated impact of income inequality on growth is very similar to that of the baseline sample of 77 countries. Also, the significance of the effect is maintained. In general, all the robustness tests provide broad support to the results obtained in the baseline empirical model.

[24] Barro (2000) argues that the relationship between net Gini and economic growth may be affected by the redistribution efforts, which may act as a disincentive for investment and economic growth by creating more distortions.

[25] While the inclusion of all the potential covariates in the same specification may be desirable, this would create challenges due to the large number of parameters to be estimated with relatively short time series. Also, other variables could be considered as potential controls, however these are often not available for a sufficient and consecutive number of years.

[26] Our robustness tests include the variables commonly used in the literature. However, to the extent that other variables have an impact on real per capita GDP growth beyond the one coming from income inequality, results could still suffer from omitted variable bias.

Figure 9: Robustness Tests for the Composite Response of Real per Capita GDP Growth to Income Inequality Growth



Source: Authors' calculations.
Notes: Confidence intervals are calculated from a resampling simulation with 500 repetitions. The panels report the response of real per capita GDP growth to a shock in income inequality growth from a PSVAR with the following variables and orderings: real per capita GDP growth and market Gini (top left); net Gini and real per capita GDP growth (top center); redistribution, real per capita GDP growth, and net Gini (top right); real per capita GDP growth, real investment growth, and net Gini (bottom left); terms of trade, real per capita GDP growth, and net Gini (bottom center); real per capita GDP growth and net Gini including data for countries with stationary series (bottom right).

5 Conclusions

Despite the relevance of the topic, the literature on the relationship between inequality and growth is far from being conclusive. From the theoretical point of view, income inequality can either propel growth by offering incentives to people to be more productive or impair it by reducing demand and worsening health and education for the poorer, among other reasons. At the same time, the empirical results which often uncover a negative effect of income inequality on growth are plagued with econometric issues. This paper reassesses the impact of income inequality on economic growth by using the largest dataset available on income inequality, allowing for reverse causality and dynamics, correcting for cross-section dependence, and relaxing the assumption of cross-country homogeneity.

Our findings suggest that income inequality is converging across countries, and that its impact on economic growth is heterogeneous. More specifically, countries are generally converging toward similar levels of income inequality, with Europe, the Middle East and Central Asia, and advanced economies experiencing within convergence. Employing a heterogeneous PSVAR, we find that the median response of real per capita GDP growth to shocks in income inequality is negative and significant for the full sample. An in-depth look, however, reveals that the effect is dispersed and that at least one fourth of the countries in the sample presents a positive effect. The only country groups for which we find evidence of a significant negative effect are the Middle East and Central Asia, the Western Hemisphere, and emerging markets.

We also find that institutional quality is correlated with the strength of the relationship between inequality and growth. In particular, we find evidence that a stronger control of corruption, rule of law, and government effectiveness can reduce the negative effect of income inequality on economic growth. Possibly, these results are linked to the fact that governments characterized by high institutional quality generally provide a more transparent framework to compete and increase productivity, hence attenuating the deleterious effect of income inequality on economic activity.

Looking forward, other challenges remain to further analyze the relationship between income and its distribution. In particular, as time goes by, the availability of more observations will permit building a more comprehensive empirical model, including all possible covariates at the same time. Moreover, the relationship between income inequality and growth is likely to be non-linear and country-specific non-linearities should be incorporated in such model.

References

Alesina, Alberto and Roberto Perotti (1996). "Income Distribution, Political Instability, and Investment". In: *European Economic Review* 40.6, pp. 1203–1228.

Alesina, Alberto and Dani Rodrik (1991). *Distributive Politics and Economic Growth*. Tech. rep. National Bureau of Economic Research.

Atems, Bebonchu and Jason Jones (2015). "Income Inequality and Economic Growth: a Panel VAR Approach". In: *Empirical Economics* 48.4, pp. 1541–1561.

Banerjee, Abhijit V and Esther Duflo (2003). "Inequality and Growth: What Can the Data Say?" In: *Journal of Economic Growth* 8.3, pp. 267–299.

Barro, Robert J (2000). "Inequality and Growth in a Panel of Countries". In: *Journal of Economic Growth* 5.1, pp. 5–32.

Batabyal, Sourav and Abdur Chowdhury (2015). "Curbing Corruption, Financial Development and Income Inequality". In: *Progress in Development Studies* 15.1, pp. 49–72.

Bazzi, Samuel and Michael A Clemens (2013). "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth". In: *American Economic Journal: Macroeconomics* 5.2, pp. 152–186.

Benabou, Roland (1996). "Inequality and Growth". In: *NBER Macroeconomics Annual 1996, Volume 11*. MIT Press, pp. 11–92.

— (2000). "Unequal Societies: Income Distribution and the Social Contract". In: *American Economic Review*, pp. 96–129.

Benhabib, Jess and Mark M Spiegel (2000). "The Role of Financial Development in Growth and Investment". In: *Journal of Economic Growth* 5.4, pp. 341–360.

Castelló-Climent, Amparo (2010). "Inequality and Growth in Advanced Economies: an Empirical Investigation". In: *The Journal of Economic Inequality* 8.3, pp. 293–321.

Chotikapanich, Duangkamon, William E Griffiths, DS Prasada Rao, and Vicar Valencia (2012). "Global Income Distributions and Inequality, 1993 and 2000: Incorporating Country-Level Inequality Modeled with Beta Distributions". In: *Review of Economics and Statistics* 94.1, pp. 52–73.

Chudik, Alexander and M Hashem Pesaran (2013). "Large Panel Data Models with Cross-Sectional Dependence: a Survey". In: *CAFE Research Paper* 13.15.

Dabla-Norris, Era, Kalpana Kochhar, Nujin Suphaphiphat, Frantisek Ricka, and Evridiki Tsounta (2015). "Causes and Consequences of Income Inequality: a Global Perspective". In: International Monetary Fund SDN/15/13.

De Dominicis, Laura, Henri LF De Groot, and Raymond JGM Florax (2006). "Growth and Inequality: a Meta-Analysis". In: *Tinbergen Institute Discussion Paper* 064/3.

Deininger, Klaus and Lyn Squire (1996). "A New Data Set Measuring Income Inequality". In: *The World Bank Economic Review* 10.3, pp. 565–591.

Easterly, William (2007). "Inequality Does Cause Underdevelopment: Insights From a New Instrument". In: *Journal of Development Economics* 84.2, pp. 755–776.

Evans, Paul (1998). "Using Panel Data to Evaluate Growth Theories". In: *International Economic Review*, pp. 295–306.

Forbes, Kristin J (2000). "A Reassessment of the Relationship Between Inequality and Growth". In: *American Economic Review*, pp. 869–887.

Galor, Oded and Daniel Tsiddon (1997). "Technological Progress, Mobility, and Economic Growth". In: *American Economic Review*, pp. 363–382.

Glaeser, Edward, Jose Scheinkman, and Andrei Shleifer (2003). "The Injustice of Inequality". In: *Journal of Monetary Economics* 50.1, pp. 199–222.

Góes, Carlos (2016). “Testing Piketty’s Hypothesis on the Drivers of Income Inequality: Evidence from Panel VARs with Heterogeneous Dynamics”. In: International Monetary Fund Working Paper 16/160.

Halter, Daniel, Manuel Oechslin, and Josef Zweimüller (2014). “Inequality and Growth: the Neglected Time Dimension”. In: *Journal of Economic Growth* 19.1, pp. 81–104.

Im, Kyung So, M Hashem Pesaran, and Yongcheol Shin (2003). “Testing for Unit Roots in Heterogeneous Panels”. In: *Journal of Econometrics* 115.1, pp. 53–74.

Iversen, Torben and John D Stephens (2008). “Partisan Politics, the Welfare State, and Three Worlds of Human Capital Formation”. In: *Comparative Political Studies* 41.4-5, pp. 600–637.

Jenkins, Stephen P (2015). “World Income Inequality Databases: an Assessment of WIID and SWIID”. In: *The Journal of Economic Inequality* 13.4, pp. 629–671.

Kaldor, Nicholas (1957). “A Model of Economic Growth”. In: *The Economic Journal* 67.268, pp. 591–624.

Kao, Chihwa (1999). “Spurious Regression and Residual-Based Tests for Cointegration in Panel Data”. In: *Journal of Econometrics* 90.1, pp. 1–44.

Kraay, Aart (2015). “Weak Instruments in Growth Regressions: Implications for Recent Cross-Country Evidence on Inequality and Growth”. In: World Bank Policy Research Working Paper 7494.

Kuznets, Simon (1955). “Economic Growth and Income Inequality”. In: *American Economic Review* 45.1, pp. 1–28.

Lazear, E and S Rosen (1981). “Rank-Order Tournaments as Optimum Labor Contracts”. In: *Journal of Political Economy* 89.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu (2002). “Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties”. In: *Journal of Econometrics* 108.1, pp. 1–24.

Li, Hongyi and Heng-fu Zou (1998). “Income Inequality is Not Harmful for Growth: Theory and Evidence”. In: *Review of Development Economics* 2.3, pp. 318–334.

Maddala, Gangadharrao S and Shaowen Wu (1999). “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test”. In: *Oxford Bulletin of Economics and Statistics* 61.S1, pp. 631–652.

Milanovic, Branko (2013). “Global Income Inequality in Numbers: In History and Now”. In: *Global policy* 4.2, pp. 198–208.

Mishra, Prachi, Peter Montiel, Peter Pedroni, and Antonio Spilimbergo (2014). “Monetary Policy and Bank Lending Rates in Low-Income Countries: Heterogeneous Panel Estimates”. In: *Journal of Development Economics* 111, pp. 117–131.

Morgan, Jana and Nathan J Kelly (2013). “Market Inequality and Redistribution in Latin America and the Caribbean”. In: *The Journal of Politics* 75.03, pp. 672–685.

Okun, Arthur M (2015). *Equality and Efficiency: The Big tradeoff*. Brookings Institution Press.

Ostry, Jonathan David and Andrew Berg (2011). “Inequality and Unsustainable Growth: Two Sides of the Same Coin?” In: International Monetary Fund SDN/11/08.

Ostry, Jonathan David, Andrew Berg, and Charalambos G Tsangarides (2014). “Redistribution, Inequality, and Growth”. In: International Monetary Fund SDN/14/02.

Pedroni, Peter (1999). “Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors”. In: *Oxford Bulletin of Economics and statistics* 61.s 1, pp. 653–670.

— (2004). “Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis”. In: *Econometric Theory* 20.03, pp. 597–625.

— (2013). “Structural Panel VARs”. In: *Econometrics* 1.2, pp. 180–206.

Pedroni, Peter and James Yudong Yao (2006). “Regional Income Divergence in China”. In: *Journal of Asian Economics* 17.2, pp. 294–315.

Persson, Torsten and Guido Tabellini (1992). “Growth, Distribution and Politics”. In: *European Economic Review* 36.2-3, pp. 593–602.
— (1994). “Is Inequality Harmful for Growth?” In: *American Economic Review*, pp. 600–621.
Pesaran, M Hashem and Ron Smith (1995). “Estimating Long-Run Relationships from Dynamic Heterogeneous Panels”. In: *Journal of Econometrics* 68.1, pp. 79–113.
Phillips, Peter CB and Donggyu Sul (2007). “Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence”. In: *Journal of Econometrics* 137.1, pp. 162–188.
Rajan, Raghuram G (2011). *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton University Press.
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.
Saint-Paul, Gilles and Thierry Verdier (1993). “Education, Democracy and Growth”. In: *Journal of Development Economics* 42.2, pp. 399–407.
Sala-i-Martin, Xavier (2006). “The World Distribution of Income: Falling Poverty and... Convergence, Period”. In: *The Quarterly Journal of Economics*, pp. 351–397.
Solt, Frederick (2016). “The Standardized World Income Inequality Database”. In: *Social Science Quarterly* 97.
Stiglitz, Joseph E (2012). *The Price of Inequality: How Today’s Divided Society Endangers Our Future*. WW Norton & Company.
Yepes, Concepción Verdugo, Peter L Pedroni, and Xingwei Hu (2015). “Crime and the Economy in Mexican States: Heterogeneous Panel Estimates (1993-2012)”. In: International Monetary Fund Working Paper 15/121.

Appendix A. Country Groups and Data Sources

We create regional groups according to the IMF regional departments and we classify countries in income levels according to the April 2016 IMF World Economic Outlook. Specifically, the regional groups include the following countries:

Africa: Botswana, Ghana, Malawi, Mali, Nigeria, Rwanda, Sierra Leone, South Africa, Uganda, Zambia, Zimbabwe.
Asia and Pacific: Australia, Bangladesh, China, India, Indonesia, Japan, Malaysia, New Zealand, Philippines, Singapore, Sri Lanka, Taiwan.
Europe: Austria, Belarus, Belgium, Bulgaria, Croatia, Cyprus, Czech Rep., Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Lithuania, Luxembourg, Norway, Poland, Portugal, Russian Federation, Slovakia, Sweden, Switzerland, Turkey, Ukraine, United Kingdom.
Middle East and Central Asia: Armenia, Egypt, Jordan, Kazakhstan, Mauritania, Morocco, Pakistan, Tajikistan, Tunisia.
Western Hemisphere: Argentina, Brazil, Canada, Chile, Colombia, Costa Rica, Dominican Rep., Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay, Peru, United States.

And the income levels include the following countries:

Advanced Economies: Australia, Austria, Belgium, Canada, Cyprus, Czech Rep., Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Luxembourg, New Zealand, Norway, Portugal, Singapore, Slovakia, Sweden, Switzerland, Taiwan, United Kingdom, United States.
Emerging Markets: Argentina, Armenia, Belarus, Botswana, Brazil, Bulgaria, Chile, China, Colombia, Costa Rica, Croatia, Dominican Rep., Ecuador, Egypt, El Salvador, Guatemala, Hungary, India, Indonesia, Jordan, Kazakhstan, Malaysia, Mexico, Morocco, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Russian Federation, South Africa, Sri Lanka, Tunisia, Turkey, Ukraine.
Low Income Developing Countries: Bangladesh, Ghana, Honduras, Malawi, Mali, Mauritania, Nigeria, Rwanda, Sierra Leone, Tajikistan, Uganda, Zambia, Zimbabwe.

Table A.1 lists all variables used in the paper, along with the source and the scale:

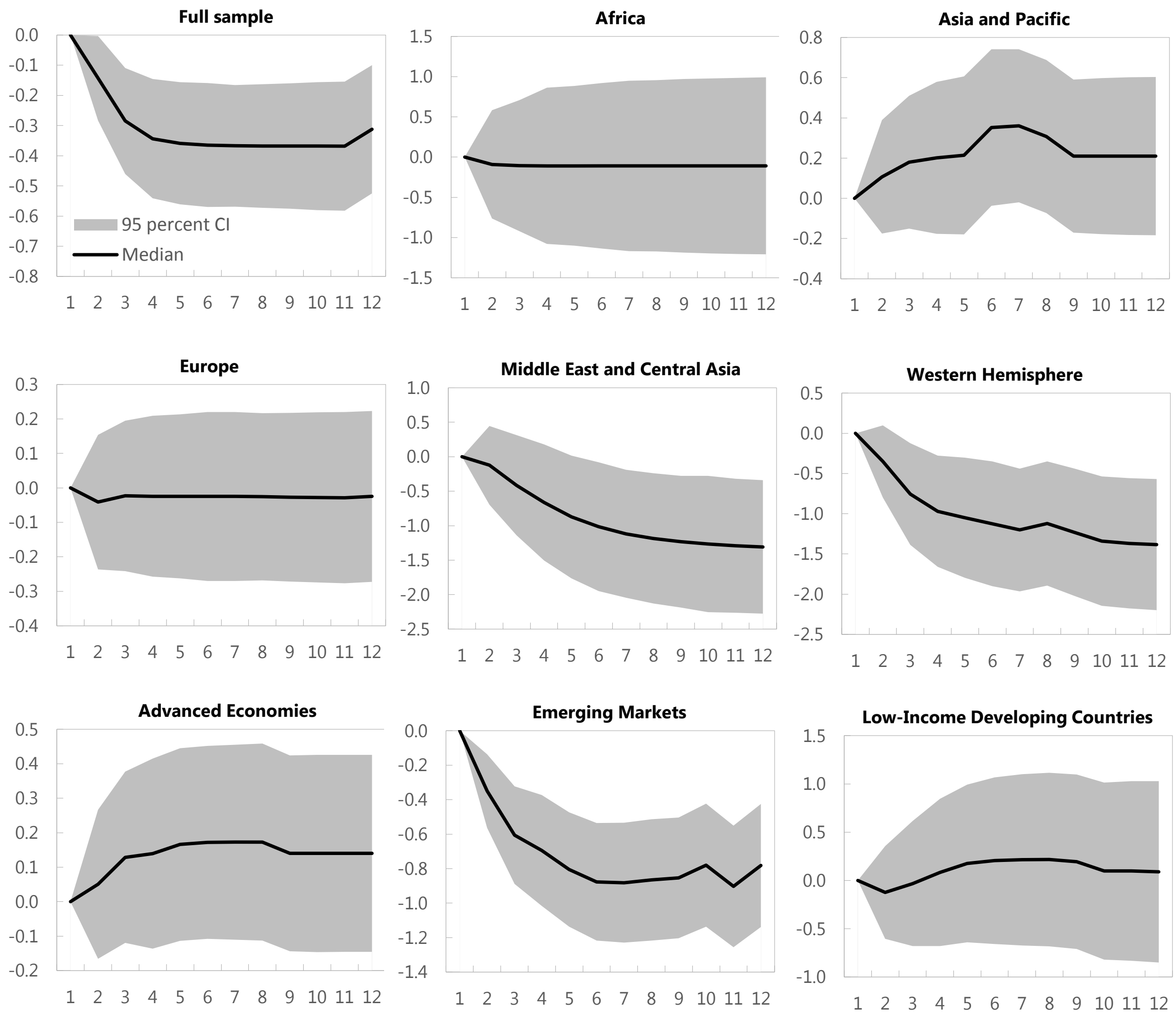
Table A.1: Data Sources

Variable	Source	Scale
Net Gini	SWIID 5.1	Index
Market Gini	SWIID 5.1	Index
Redistribution	SWIID 5.1	Percent
Real per capita GDP	PWT 9.0	PPP
Control of corruption	World Governance Indicators	Index
Rule of law	World Governance Indicators	Index
Government effectiveness	World Governance Indicators	Index
Political stability	World Governance Indicators	Index
Economic Freedom of the World Index	Fraser Institute	Index
Years of schooling	Barro Lee Database	Units
Education expenditure	IMF's Fiscal Affairs Department database	Percent of GDP
Real investment growth	World Economic Outlook	Percent
Terms of trade	World Economic Outlook	Index

Appendix B. Accumulated Impulse Response Functions

Figure B.1 reports the median accumulated composite responses of real per capita GDP growth to income inequality growth for the full sample as well as sample subsets.

Figure B.1: Accumulated Composite Response of Real per Capita GDP Growth to Income Inequality Growth by Region and Income Level and Confidence Intervals



Notes: Confidence intervals are calculated from a resampling simulation with 500 repetitions.
Source: Authors' calculations.