Oechslin, Manuel; Halter, Daniel; Zweimüller, Josef

Conference Paper

Inequality and Growth: The Neglected Time Dimension

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Empirical Studies of Inequality, No. E4-V3

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Oechslin, Manuel; Halter, Daniel; Zweimüller, Josef (2010) : Inequality and Growth: The Neglected Time Dimension, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Empirical Studies of Inequality, No. E4-V3, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/37461

Inequality and Growth: The Neglected Time Dimension

Daniel Halter, University of Zurich
Manuel Oechslin, Tilburg University*
Josef Zweimüller, University of Zurich

July 31, 2009

Abstract

When it comes to the inequality-growth relationship, the empirical literature offers contradictory assessments: Estimators based on time-series variation only (i.e., differences-based estimators) indicate a strong positive link while estimators also exploiting the cross-sectional variation (i.e., level-based estimators) suggest a negative relationship. Taking advantage of a new dataset, the present paper confirms this conflicting pattern – but also offers an explanation on the basis of a simple model: The seemingly inconsistent empirical results just mirror different aspects of reality. More specifically, we argue that the differences-based methods are prone to reflect the (mostly) positive *short-* or *medium-run* implications of inequality while the level-based estimators also incorporate the negative consequences – which require more time to materialize. Thus, the latter estimates come close to reflect the adverse overall impact of inequality in the *long run*.

JEL classification: O11; O15; O43; C23

Keywords: Inequality, growth, medium-run effects, long-run effects

*Corresponding author. Contact details: Tilburg University, Department of Economics, PO Box 90153, 5000 LE Tilburg, The Netherlands; phone: +31 13 4663650; fax: +31 13 4663042, email: m.oechslin@uvt.nl.

1 Introduction

The impact of inequality on economic performance is a much debated topic in economics. During the past two decades, theoretical work has come up with a substantial number of channels through which inequality may influence economic growth either in a *positive* or *negative* direction (see, e.g., Galor, 2009, for a recent and comprehensive overview). These theoretical contributions have made abundantly clear that the impact of inequality is a complex one. Its influence is likely to depend on, among other things, the specifics of a country (e.g., the stage of economic development; the extent of market failures; the form of government) or the time horizon considered (e.g., medium run vs. long run) and thus cannot be expected to be clear-cut in reality. Not surprisingly thus, this theoretical ambiguity is mirrored in the empirical literature which – mainly based on broad panels of countries – finds both significantly positive and negative effects, and sometimes no effects at all.

Looking more closely at the empirical picture, however, reveals a certain pattern in the large body of existing results: Estimates which are based on time-series variation only (e.g., estimations relying on first-difference estimators as those in Forbes, 2000) find a strong positive impact of inequality; yet, estimations which also exploit the cross-sectional variation in the data (e.g., estimates based on random-effects estimators as in Barro, 2000) tend to find negative relationships, and significantly so in samples of less-advanced economies.[1]

Against this backdrop, the present paper deals with the inequality-growth relationship both from an empirical and a theoretical perspective. The paper makes two contributions: *First*, it takes advantage of a new and more comprehensive inequality data set and confirms the seemingly contradictory findings in the literature: Based on a much larger number of observations, the first-difference GMM estimator still consistently indicates a strong positive relation the while system GMM estimator (which also exploits the cross-sectional variation) continues to find a negative link (and significantly so in all but the richest countries).

The *second* contribution is to interpret these results through the lens of the existing theory. We will argue that both results are reasonable and just represent different aspects of reality. Specifically, we suggest that the standard regression equation (which imposes a monotonic impact of inequality on growth) is miss-specified in a way that induces (i) the first-difference GMM estimators to systematically pick up the positive channels; (ii) the system GMM estimator to primarily reflect the negative ones. Besides an informal argumentation, we also

[1] Earlier studies which find a positive relation on the basis of time-series variation only include, among others, Li and Zou (1998) and Benhabib and Spiegel (1997). Earlier work documenting a negative link on the basis of cross-sectional variation consists of, among others, Alesina and Rodrik (1994) and Persson and Tabellini (1994).

introduce a parsimonious model to convey our explanation in a more precise way.

Our argumentation relies on one central observation: In our reading, the theoretical literature clearly suggests that the positive and negative effects of inequality cluster in a very specific way.[2] Most of the positive effects (e.g., those operating through convex savings functions or through incentives for innovative activity) rely on purely economic mechanisms and thus tend to materialize relatively fast, arguably in the *short* or *medium run*. Most of the negative effects, however, involve the political process or require the rise of socio-political movements (or changes in the composition of educational attainments in the population); thus, it is reasonable to argue that the negative effects are primarily *long-run* effects.

Given this specific pattern, the seemingly contradictory empirical results can be interpreted in a natural way. Intuitively, the first-difference GMM estimator regresses changes in output on only slightly lagged changes in inequality. As a result, if – for instance – inequality goes up, the positive short- or medium-run effects are associated with positive changes in inequality but – due to the specific time structure of the panel – the subsequent negative changes (i.e., those coming from the long-run effects) are not systematically attributed to the very same initial increase of the Gini coefficient (but just treated as noise). So the first-difference estimates are likely to be biased in the sense that they only reflect the benign short- or medium-run effects but leave out the potentially severe long-run consequences. Panel *a.* of Figure 1 gives a first illustration of this positive short- or medium-run relationship in our data.

Figure 1 here

The system GMM estimator is nonetheless likely to find a negative relationship, in particular if the negative long-run effects are strong and thus dominate the positive short- or medium-run effects. The argument is that, since inequality is very persistent, only few countries in a given sample are actually waiting for the powerful long-run effects to set in. Put differently, if the negative effects dominate, a large fraction of the observations is either of the type "low level of inequality and high level of output" or "high level of inequality and low level of output." Thus, since the system GMM estimator also exploits cross-country variation (and hence also relies on the regression equation in levels), it is prone to reflect a negative relationship (Panel

[2] As we will discuss in Section 3, inequality can promote growth by, e.g., fostering aggregate savings (Kuznets, 1955; Kaldor, 1955); by promoting the realization of high-return projects (Rosenzweig and Binswanger, 1993); by stimulating R&D (Foellmi and Zweimueller, 2006). On the other hand, inequality may hamper growth by, e.g., promoting expensive fiscal policies (Perotti, 1993); by inducing an inefficient state bureaucracy (Acemoglu et al. 2008); by hampering human capital formation (Galor and Zeira, 1993; Galor and Moav, 2004); by leading to political instability (Bénabou, 1996); by undermining the legal system (Glaeser et al., 2003).

b. of Figure 1) – which corresponds to the true overall impact of inequality.

This paper is part of a small literature which tries to get a better grasp of the empirical picture with respect to inequality and growth. Earlier contributions include, among others, Banerjee and Duflo (2003) and Voitchovsky (2005). The former paper presents evidence suggesting that changes in inequality (in any direction) are associated with reduced growth in the short run; as a result, the standard regression equation might be miss-specified in a way that – misleadingly – makes differences-based estimators indicate a positive relationship. Voitchovsky (2005), by contrast, argues that inequality coming from the top end of the distribution is indeed likely to promote growth while bottom-end inequality tends to be harmful. She thus suggest controlling separately for inequality coming from different parts of the distribution (and finds supportive evidence in a panel of rich countries). None of these papers, however, focuses specifically on the time dimension, and so we view our paper as complementary.

The remainder of this paper is organized as follows. Section 2 presents the empirical results and links them to the earlier literature. In Section 3, we interpret our findings, first on the basis of an informal argumentation and then with the help of a simple model. Section 4 concludes.

2 Empirical Analysis

2.1 Specification and Estimation

Specification and data. We rely on standard 5-year panel data model which is similar to those used in several recent empirical studies on growth (e.g., Caselli et al., 1996; Barro, 2000; Forbes, 2000; Voitchovsky, 2005). Specifically, we estimate the linear regression equation

$$y_{it} - y_{it-1} = \gamma y_{it-1} + \mathbf{x}'_{it}\boldsymbol{\delta} + \zeta_t + \eta_i + v_{it}, \tag{1}$$

where $i = 1, \cdots, N$ denotes a particular country and $t = 1, \cdots, T$ is time (with t and $t-1$ five years apart). The variable y stands for the log of real GDP per capita so that the left-hand side of equation (1) approximately represents country i's five-years growth between $t-1$ and t. On the right-hand side, we have y_{it-1} to control for convergence; a vector $\mathbf{x}_{it}$ consisting of variable (and observable) country characteristics; a period-specific effect ζ_t to capture productivity changes common to all countries; a country-specific effect η_i to capture time-invariant (and unobserved) country characteristics; an idiosyncratic error term v_{it}.

The vector $\mathbf{x}_{it}$ consists of the Gini index and three additional standard control variables. Following the recent literature (e.g., Perotti, 1996; Forbes, 2000), these additional variables are the average years of secondary schooling in the population above 25 (separately for males

and females) and the price level of investment (to control for market distortions). In general, the explanatory variables are measured at the beginning of each 5-year period. In case of inequality, this is not always possible because the Gini index is rarely available on an annual basis. In these cases, we take the last available value in the previous 5-year period.

The analysis includes up to 90 countries and covers the period from 1966 to 2005 (and thus includes eight 5-year periods). The GDP per capita data comes from the World Development Indicators (WDI; World Bank, 2006) and is in constant 2000 US$. The Deininger and Squire (1996) data base serves as the primary source for the inequality data. However, in order to broaden our sample in the cross-sectional as well as the time-series dimension, we also rely on a subsidiary source, the UNU-WIDER (2008) data base.[3] Finally, the education data comes from Barro and Lee (2000) and the source for the price of investment is Heston et al. (2006; PWT 6.2). More detailed sources and definitions for these variables as well as some summary statistics are presented in Table 1.

Table 1 *here*

Estimation methods. It has been extensively discussed in the literature that the standard panel data methods (i.e., fixed-effects [FE] and random-effects [RE] estimations) are unlikely to provide consistent estimates of the parameters γ and $\boldsymbol{\delta}$ (see, e.g., Bond et al., 2001). Obviously, using the random-effects estimator is problematic because the unobserved country effect, η_i, is most likely correlated with the other explanatory variables. A second problem becomes immediately transparent when we rewrite model (1) as

$$y_{it} = (\gamma + 1)y_{it-1} + \mathbf{x}'_{it}\boldsymbol{\delta} + \zeta_t + \eta_i + v_{it}. \tag{2}$$

As already discussed in previous work, equation (2) highlights that controlling for convergence in a panel data growth model actually introduces a lagged dependent variable. As a result, even if equations (1) and (2) give an accurate description of reality, both the RE estimator and the FE estimator are likely to give inconsistent estimates of γ and $\boldsymbol{\delta}$.

To deal with these problems, the literature has developed specific GMM estimation techniques, most notably the first-difference GMM estimator and the system GMM estimator. The first-difference GMM estimator was developed by Arellano and Bond (1991) and is similar to the FE estimator in the sense that it employs only *within-country* variation. More specifically, the idea is to eliminate the country-specific effect by differencing model (2) and then to use

[3] Note that some of the Gini coefficients are based on income while others are based on expenditures (which, in general, are more equally distributed). To account for this inconsistency, we follow Deininger and Squire's (1996) and Forbes' (2000) suggestion to add 6.6 points to expenditure-based coefficients.

sufficiently lagged values of y and $\mathbf{x}$ as instruments. This procedure generates moment conditions of the form $E\{\triangle v_{it} y_{it-s}\} = 0$ (with $t \geq 3$ and $s \geq 2$) and $E\{\triangle v_{it} \mathbf{x}_{it-s}\} = \mathbf{0}$ (with $t \geq 3$ and $s \geq 1$) which can be used to obtain consistent estimates of γ and $\boldsymbol{\delta}$ if $N \to \infty$ (provided that the empirical model is true and the standard assumptions are satisfied).[4]

However, although the first-difference GMM estimator "solves" the problems of unobserved heterogeneity and lagged dependent variables, it has been criticized for the fact that it does not exploit the variation in levels across countries. The main concern is that the cross-sectional variation embodies a large part of the available information since within-country inequality is quite persistent.[5] Thus, ignoring this cross-sectional variation may give rise to unnecessarily large biases and imprecision. Yet, this problem can be mitigated if the first differences of y_{it} and $\mathbf{x}_{it}$ are uncorrelated with the country-specific effect, η_i. Under these circumstances, additional moment conditions can be obtained on the basis of the level-form of equation (2), with the first differences as instruments: $E\{(\eta_i + v_{it}) \triangle y_{it-1}\} = 0$ (with $t \geq 3$) and $E\{(\eta_i + v_{it}) \triangle \mathbf{x}_{it-s}\} = \mathbf{0}$ (with $t \geq 3$ and $s \in \{0, 1\}$).[6] Combining these latter moment conditions with those in first differences allows then for computing the so-called system GMM estimator which was pioneered by Arellano and Bover (1995) and Blundell and Bond (1998). In comparison to the first-difference approach, the system GMM procedure does better in terms of efficiency since – like the RE estimator – it also exploits the *cross-country* variation in the data.

In what follows, we will apply both GMM estimation techniques to our new dataset and document that – consistent with the earlier empirical literature – the two approaches lead to systematically different estimation results.

2.2 Results

Time-series variation only. We now go through the first-difference estimation results. To connect with the previous literature, we first present evidence based on a sample which is similar to that in Forbes (2000) in terms of countries included and periods covered. We then show that these results are quite robust to the inclusion of additional countries and more recent observations (from the WIDER data base) as well as to a number of other modifications.

The first column of Table 2 gives the results based on the Forbes sample (which includes

[4] See, e.g., Bond et al. (2001) for the details. Moreover, note that the timing restriction in the second set of conditions applies only if $\mathbf{x}_{it}$ is predetermined (i.e., uncorrelated with v_{it}). If $\mathbf{x}_{it}$ were endogenous, the instrument had to be lagged by at least by two periods ($s \geq 2$).

[5] This observation also applies to our dataset: The adjusted R^2 from a regression of the Gini coefficient on country dummies is 0.84 (and rises only to 0.85 if time dummies are also included).

[6] The timing restriction in the second set of conditions again applies if $\mathbf{x}_{it}$ is predetermined.

42 countries and covers the 1965-1995 period). As Forbes, we find a significant positive impact of inequality on growth, and the magnitude of the effect is very similar: On an annualized basis, our estimates imply a coefficient of 0.0015 while Forbes (2000) reports one of 0.0013. As the second column shows, the coefficient on inequality remains significant and comparable in size after extending the sample by two additional 5-year periods (i.e., the 1996-2000 and 2001-2005 periods). Similarly, as documented in the third column, the inclusion of 28 additional countries does not change the basic empirical finding: Higher inequality has a significantly positive impact on (short-run) growth, albeit the effect is somewhat smaller in the broader country sample (which includes a larger fraction of less-advanced countries).[7]

Table 2 *here*

The remaining columns of Table 2 document the robustness of this empirical outcome to some natural variations. First, the estimates in columns (4) and (5) are based on subsets of the full sample. Specifically, column (4) shows the impact of inequality in countries which are classified as high income or upper-middle income (according to the definition by the World Bank, 2009); column (5) provides the corresponding results for the remaining countries (lower-middle income or low income). Apparently, although the two subsets contain very different economies, the estimated impact of inequality is still significantly positive in both cases and of very similar size across the two country groups.

The second modification concerns the time structure of the panel. In order to check whether the above results are not just an artifact of the 5-year structure, the estimates in columns (6) and (7) are based on four 10-year periods. The results suggest that higher inequality tends to foster growth also over this medium time horizon, and the size of the estimated impact is somewhat larger: For instance, on an annualized basis, the coefficient in the forth column (5-year periods; high and up-mid countries) is 0.00082 while the corresponding coefficient for the 10-year structure is 0.00114. However, the estimates are less precise – which is not surprising given that we have a much smaller number of observations.

Note, finally, that the validity of the first-difference GMM estimator depends on the absence of serial correlation in the error terms, v_{it}. This means that the differenced error terms should not show second-order serial correlation (though they have a first-order correlation by construction). The statistics $M1$ and $M2$ in Table 2 give the t-values associated with the tests for, respectively, first-order and second-order correlation in the $\triangle v_{it}-$series. As the numbers

[7] Note that 20 out of these 28 additional countries are low income countries or lower-middle income countries according to the classification by the World Bank (2006). As a result, in the full sample, 47% of the countries fall into these two categories (while the rest belongs to the categories upper-middle income or high income).

show, serial correlation may only be an issue in the very first regression (Forbes replication) but not in columns (2) – (7).

Time-series and cross-sectional variation. Table 3 presents the results based on the system GMM estimator. The first column presents the estimates based on the full sample. Unlike in all the regression shown in the previous table, the estimated impact of inequality on growth is now negative, yet not significantly so.[8] More precise results can be gained, however, by splitting the country sample again along income classes (columns 2 – 4). It turns out that, as shown in the second column, also the system GMM estimates indicate a positive impact of inequality among the small group of high-income countries. However, there is no significant relationship among upper-middle-income countries (third column),[9] and – most importantly – the system GMM estimates indicate a negative impact in the large group of countries with lower-middle income or low income (fourth column). Note further that, as it is the case above, switching to a 10-year panel structure confirms the results received under the standard 5-year structure (columns 6 and 7 of Table 3).

Table 3 *here*

So, even though the test statistics at the bottom of Table 3 support the the validity of the instruments also with this estimation strategy, the system GMM approach paints a decidedly different picture than first-difference estimator: While the latter uniformly points to a positive relationship (and thus confirms the results by, e.g., Li and Zou, 1998; Forbes, 2000), the findings here suggest that the impact of inequality on growth is negative (or at least non-positive) in countries which are not among the richest. Note that this result is perfectly in line with Barro's (2000) random-effects analysis (which also exploits cross-sectional as well as time-series variation) and also matches with the results in earlier OLS-based studies such as Alesina and Rodrik (1994) or Persson and Tabellini (1994).

3 Interpreting the Empirical Results

The present section looks at how these seemingly contradictory estimation results could be interpreted and reconciled. Subsection 3.1 takes an informal approach and proceeds in two steps.

[8] The number of countries included in the sample rises to 90 since the system GMM estimator also includes moment conditions on the basis of the level form of the regression equation (and hence - in contrast to the first-difference estimator - does not strictly require two consecutive observations).

[9] If we combine – as in Table 2 – high-income countries and upper-middle income countries in one sample, the estimated coefficient on inequality is insignificant (result not reported in the table).

The first step is to stress that, in fact, the existing literature suggest that both relationships should be present in reality. In the second step, we then argue that regression equation (1) is miss-specified so that the two different GMM estimators (i.e., the first-difference estimator and the system estimator) are prone to systematically reflect just one of the two relationships, namely the positive one in case of the differences-based approach and the negative one if the estimator also exploits cross-sectional variation. In Subsection 3.2, we continue by introducing a simple model to make the informal argumentation more precise.

3.1 Broad Arguments

Short-run and medium-run effects vs. long-run effects. Inequality affects economic performance through many channels, and the theoretical literature prominently discusses both negative as well as positive effects. As for the positive channels, the literature has argued since long that savings functions tend to be convex in wealth (see, e.g., Kuznets, 1955; Kaldor, 1955). So, other things equal, higher inequality is associated with higher aggregate savings and thus faster convergence to the balanced growth path. More recently, the focus has been on the impact of inequality on the selection of physical investment projects (see, e.g., Matsuyama, 2000, in particular Section 4). The main argument here is that, if the financial system is imperfect, access to external finance depends on personal wealth. As a result, if wealth is widely spread among the population, nobody may be able to raise sufficient funds to realize high-return projects which require large minimum investments. In this case, a more concentrated distribution of productive assets may put at least a limited number of entrepreneurs into a position to realize such projects – and thus boosts growth.[10] This effect is reinforced by the fact that the high-return projects are often the more risky ones (see, e.g., Rosenzweig and Binswanger, 1993). As a result, with a relatively equal wealth distribution, the number of entrepreneurs who are sufficiently rich to absorb significant risks may be very limited. So, once again, a more concentrated distribution of wealth may multiply the number of realized high-return projects. Finally, the literature also discusses positive demand-side effects. With a more unequal distribution, a larger fraction of total demand falls on "high-end" products (as opposed to goods satisfying basic needs). Thus, potential innovators face larger home markets which more easily support the investments required to develop novel or better varieties (see, e.g., Foellmi and Zweimueller, 2006). Clearly, this positive demand-side effect is more relevant

[10] It has also been argued that, with convex technologies and financial markets imperfections, higher inequality deteriorates economic performance because investment returns are more heterogeneous. However, as shown by Foellmi and Oechslin (2008), this is by no means a robust theoretical prediction.

in advanced economies where R&D is the main driver of growth.

While all these positive effects work through different channels, they have one thing in common: They all emphasize purely economic mechanisms. As a result, we should expect these effects to materialize relatively fast. This, however, is clearly different in case of the negative channels. Some of the most prominent negative links rely on political-economy arguments. For instance, it has been pointed out that more unequal societies tend to have higher levels of redistribution and hence higher levels of taxation – which weakens the incentives to save and invest (see, e.g., Perotti, 1993). A related argument focuses on the composition of government expenditures. With higher inequality, the decisive voter tends to supply fewer production factors (i.e., physical or human capital). As a result, he may strongly prefer direct transfers ("handouts") over productivity-enhancing investments in public goods. Finally, even if political power rest with the rich, inequality may still have a negative impact via the fiscal policy channel. As highlighted by Acemoglu et al. (2008), if inequality is high, an oligarchic government has incentives to set up an inefficient bureaucracy in order to avoid high taxation once the country is transformed into a democracy.[11] Yet, at least via these channels, changes in inequality cannot be expected to have an immediate effect on economic performance. It certainly takes time until shifts in the population's policy preferences are reflected in similar changes within the legislative body. Moreover, even with a fresh legislative in place, altering tax laws (or even changing the bureaucracy) is a time-consuming process.

Note further that also the remaining negative effects are unlikely to materialize quickly. If higher inequality reduces aggregate spending on human capital formation (see, e.g., Galor and Zeira, 1993; Galor and Moav, 2004), it arguably takes a decade or more until the effects can be felt. Similarly, it may take a long time until disaffection caused by higher inequality is bundled in social movements which then may threaten political stability (see, e.g., Bénabou, 1996) or until higher inequality has undermined the reliability of the judicial system and the security of property rights.(see, e.g., Glaeser et al., 2003).

Differences vs. levels. Our brief literature survey clearly suggests that the positive and negative effects of inequality cluster in a very specific way: The positive effects tend to materialize quickly while the negative effects need more time to emerge. The present subsection argues that it is exactly this pattern which is responsible for the different estimation results obtained above. To see this, it is convenient to look in a first step at the differences-based methods (e.g.,

[11] More generally, based on the experience of the colonization of the New World, Sokoloff and Engerman (2000) argue that huge wealth inequalities may promote institutions that protect the privileges of the elites and restrict opportunities for the broad masses – with adverse consequences for economic development.

the first-difference GMM estimator). Clearly, since these methods regress changes in output on moderately lagged changes in inequality, they are likely to pick up the short-run or medium-run effects – and thus to find a positive relationship. To give an example, if inequality goes up, aggregate output tends to respond positively in the short or medium run because, for instance, a higher wealth concentration supports a larger number of high-return investments while the supply of the public good or the quality of the institutions do not yet deteriorate. As a result, differences-based methods associate a positive change in inequality with a positive change in output but – due to the specific time structure of the panel – fail to systematically attribute the subsequent negative changes (i.e., those changes coming from the long-run effects) to the initial increase in inequality. Put differently, the negative changes are just treated as noise, and so the differences-based methods are set to find a positive effect.[12]

However, the methods which also exploit the variation in the levels (e.g., the system GMM estimator) are nonetheless likely to find a negative link, in particular if two conditions are satisfied. First, the long-run effects must dominate the short-run or medium-run effects and, second, within-country inequality has to be a rather persistent phenomenon. Note, however, that there are indeed good reasons to assume that these conditions hold. As for the relative strength of the different effects, a broad empirical literature suggest that institutional quality has a dominant impact on on economic performance (see, e.g., Acemoglu et al., 2001). Regarding persistence, our data as well as an elaborate literature support the notion that countries do not frequently undergo significant changes in inequality. To see now why under these circumstances the cross-sectional methods find a negative link, consider two countries which have shown different degrees of inequality for a while. Then, other things equal, the low-inequality country (i.e., the country with the good institutional quality) would have a higher GDP than the high-inequality country (i.e., the country with the bad institutional quality). Hence, if within-country inequality was perfectly persistent over time, the level-based methods would find a clear-cut negative link between inequality and economic performance – which is driven by the comparatively strong long-run effects. Yet, inequality is not completely persistent, and so the data-generating process creates observations which potentially "mire" the picture. For instance, following a switch from low to high inequality, we may have a number of observations with both high inequality and high output because the positive effects have already set in but the negative long-run effects are still to come. However, if within-country inequality is

[12] The argument is completely symmetric for negative changes in inequality. The time-series methods link negative changes in inequality to contemporaneous negative changes in output but fail to attribute subsequent improvements to the initial decline in inequality. Again, the long-run effect is just regarded as noise.

persistent, such transition periods are relatively rare and a large fraction of the observations is either of the type "high inequality and low output" or "low inequality and high output." Accordingly, the data points that do not fit into this latter pattern are treated as noise (i.e., driven by exogenous shocks), and the regression analysis points towards a negative relationship – which can be interpreted as the overall relationship in the long run.

3.2 A Formal Approach

We now introduce a formal model to make the above argumentation more precise. The model is highly parsimonious in the sense that it incorporates just one of the positive (i.e., short-run or medium-run) effects and one of the negative (i.e., long-run) effects. The advantage of this approach is that it conveys our argumentation in a very transparent and clean way.

Assumptions. We focus on an infinite-horizon economy which is populated by a continuum of individuals of measure 1. All agents derive utility from consumption of a single (non-storable) output good, and preferences are represented by the inter-temporal utility function

$$U_t = E_t \left\{ \sum_{s=0}^{\infty} \beta^s c_{t+s} \right\}, \tag{3}$$

where c_t denotes consumption in period t. Individuals differ regarding their endowment with the productive asset (which we may interpret as "skills," for instance). A fraction $\alpha > 1/2$ of the population (the "poor", P) is endowed with $\omega^P(D_t) < 1$ units of this asset, whereas 1 is the average endowment in the economy. The endowment of the remaining agents (the "rich", R) is then given by $\omega^R(D_t) = (1-\alpha\omega^P(D_t))/(1-\alpha) > 1$. The state variable $D_t \in \{L, H\}$ represents the degree of inequality in the economy. We assume $\omega^P(L) > \omega^P(H)$ which means that the poor have a relatively high – but still below-average – endowment in state L ("low" inequality). Note further that, at the beginning of each period, inequality may change exogenously. In particular, we have $D_t = D_{t-1}$ with probability π and $D_t \neq D_{t-1}$ with probability $1-\pi$ so that a high value of π mirrors strong *persistence* in inequality. In practice, a change in the distribution of skills may be due to a shock to the educational system which improves the quality of primary education relative to that of university education, for instance.

Regarding production side of the economy, suppose that the individuals have access to a simple linear technology of the form

$$y^i(D_t, G_t) = a^i \omega^i(D_t) X(G_t), \tag{4}$$

with $i \in \{P, R\}$, where a^i is a group-specific productivity parameter and $X(G_t)$ is the level of the public good. We further assume the rich agents to be *more productive* than the poor:

$a^R > a^P$. A natural way to think of this assumption is that the more productive technology requires a certain talent level which cannot be achieved by the poor.[13] The state variable $G_t \in \{0, 1\}$ reflects whether – in the previous period – the government has invested in the public good, with 1 indicating investment. Thus, we have $X(1) - X(0) \equiv \triangle X > 0$.

On the aggregate level, we can now easily infer that (private-sector) output is given by

$$Y(D_t, G_t) = \left(a^R - \alpha(a^R - a^P)\omega^P(D_t)\right) X(G_t). \tag{5}$$

Other things equal, Y is higher in the high-inequality state ($D_t = H$) since a larger fraction of the productive asset is allocated to the high-return technology; similarly, output is higher if the level of the public good is high ($G_t = 1$). Finally, note that $Y(L, 1) > Y(H, 0)$ if

$$\frac{X(1) - X(0)}{X(1)\omega^P(L) - X(0)\omega^P(H)} > \alpha \frac{a^R - a^P}{a^R}, \tag{6}$$

i.e., if the productivity differential between rich and poor is not too large.

Turning to the public sector, suppose that the government has access to an income stream of Z units of the final good. We can think of this income as arising from a publicly owned enterprise, the natural resource sector, etc. Regarding public spending, the government has to decide on G_{t+1} in each period t. A decision to invest is associated with a contemporaneous cost of $F < Z$ units of the final good. The budget surplus is distributed to the population in a lump-sum manner. Finally, when deciding on G_{t+1}, we assume that the government has no choice but to implement the variant preferred by the majority of the population, i.e., the poor.

An interesting equilibrium. We know show that our model is able to generate equilibrium patterns which would give rise to estimation results similar to those outlined in Section 2. The first step is to establish that the level of the public good may fluctuate over time:[14]

Proposition 1 *The equilibrium shows fluctuations in public goods provision – with positive investment in times of low inequality (i.e., $G_{t+1} = 1$ if $D_t = L$) and absent investment in times how high inequality (i.e., $G_{t+1} = 0$ if $D_t = H$) – if the following condition holds:*

$$\frac{\triangle X}{F} a^P \left(\pi\omega^P(L) + (1 - \pi)\omega^P(H)\right) \geq \frac{1}{\beta} > \frac{\triangle X}{F} a^P \left(\pi\omega^P(H) + (1 - \pi)\omega^P(L)\right). \tag{7}$$

[13] More generally, this assumption can be seen as a reduced-form representation of the notion that only relatively rich people can rely on high-return technologies because – as discussed in Subsection 3.1 – the financing of such technologies requires good access to the financial system (which the poor lack).

[14] Suppose that the social planner maximizes a utilitarian social welfare function. Then, a sufficient condition for the planner to choose $G_{t+1} = 1$ in each period t is $(\Delta X/F)a^P \geq 1/\beta$.

Proof. See Appendix. ■

Intuitively, condition (7) holds if – in case of high inequality – the poor perceive a sufficiently low stake in the production-side of the economy so that they prefer direct transfers over investment in the public good (and the other way round in case of low inequality).

We now proceed in order to characterize the relationship between changes in inequality and changes in aggregate output if the above fluctuations equilibrium is relevant:

Corollary 1 *Suppose that the condition (6) and (7) hold. Moreover, assume that inequality has been unchanged between $t-2$ and $t-1$. Then,*

(i) an increase in inequality in period t (i.e., $D_{t-1} = L \rightarrow D_t = H$) leads to a contemporaneous increase in output ($Y_{t-1} = Y(L,1) < Y_t = Y(H,1)$); however, in $t+1$, output declines sharply, with inequality either unchanged or decreasing.

(ii) a decrease in inequality in period t (i.e., $D_{t-1} = H \rightarrow D_t = L$) leads to a contemporaneous decrease in output ($Y_{t-1} = Y(H,0) > Y_t = Y(L,0)$); however, in $t+1$, output rises sharply, with inequality either unchanged or increasing.

The intuition behind Corollary 1 is that the level of the public good is a state variable and thus cannot change quickly. So an increase in inequality must lead to a positive change in output in the *short run* (i.e., with X still at the high level) but to a negative one in the *long run* (i.e., when the increase in inequality has undermined to provision of the public good). Obviously, for a decrease in inequality, the implications are exactly reverse.

Regarding the link between the levels of inequality and output we can infer the following:

Corollary 2 *Suppose that the condition (6) and (7) hold. Moreover, assume that inequality is persistent (i.e., that π is "high"). Then, over time,*

(i) a large fraction of the observations (D_t, Y_t) will either be "low" inequality and "high" output, $(L,\ Y(L,1))$, or "high" inequality and "low" output $(H, Y(H,0))$.

(ii) very few observations (D_t, Y_t) will either be "low" inequality and "(very) low" output, $(L, Y(L,0))$, or "high" inequality and "(very) high" output, $(H, Y(H,1))$.

The central point behind Corollary 2 is *persistence* in inequality. Persistence means that periods with changes in inequality – which generate observations of the type ("high" inequality/"high" output) or ("low" inequality/"low" output) – are relatively infrequent.

Estimating the relationship. We now discuss how the different estimation methods reflect the inequality-output relationship that is implied by the present model. An illustrative way

to do so is to give a graphical representation of the two corollaries in a single picture – which is done in Figure 2. To see how the figure is constructed, consider the case of an increase in inequality in period t. If the focus is on changes (Panel *a.*), the following types of observations are generated: Observation -1 in period $t-1$, observation 0 in period t (when the short-run effect on output materializes), observation 1 in period $t+1$ (when the long-run effect on output materializes but inequality is unchanged), and observation 2 in period $t+2$ (when both inequality and output are unchanged). The remaining observations in Panel *a.* are generated by going through the case of decreasing inequality. The numbers in Panel *b.* refer to the same experiment, yet from the perspective of the levels.

Figure 2 here

Figure 2 illustrates that the different aspects of the relationship between inequality and output are picked up by different estimation methods. If the relationship is assessed on the basis of changes (Panel *a.*), we can see that estimating a linear regression would give us a clear positive relationship. On the other hand, if levels are considered (Panel *b.*), fitting a linear trend line would arguably point to a significant negative impact of inequality (since observation of the types of -1, 1, and 2 are much more numerous than observations of type 1).

It might also be interesting to look at the different estimation methods from a more formal perspective. We start by deriving the formal relationship between output and inequality, given that conditions (6) and (7) hold. Taking logs on both side of equation (5) gives us

$$y_t \equiv \ln Y_t = \ln\left(1-\alpha\frac{a^R-a^P}{a^R}\omega^P(D_t)\right)+\ln\left(1+\frac{\triangle X}{X(0)}\frac{H-D_{t-1}}{H-L}\right)+\ln a^R+\ln X(0),$$

whereas the second term on the right-hand side represents the equilibrium expression for $X(G_t(D_{t-1}))$. Assume now further that $\omega^P(D_t)=1-D_t$. Then, the above expression can be approximated by the linear regression equation

$$y_{it}=\delta_1 D_{it}+\delta_2 D_{it-1}+\eta_i+v_{it}, \tag{8}$$

whereas $\delta_1 \equiv \alpha(a^R-a^P)/a^R$, $\delta_2 \equiv -\triangle X/(X(0)(H-L))$, and $\delta_1+\delta_2<0$ due to condition (5). The sum of the constant terms is represented by η (which we allow to vary across countries) and – as in equation (2) – v_{it} denotes an idiosyncratic error term which reflects exogenous influences on private-sector output.[15] Obviously, the key difference between the theory-based equation (8) and the standard equation (2) is that the former one also includes lagged inequality, D_{t-1}, while the latter one just ignores earlier levels of inequality.

[15] The constant η_i may be country-specific due to, for instance, cross-country differences in the levels of firm productivity (however, with $(a^R-a^P)/a^R$ constant across countries).

We are now able to analytically determine the biases if model (8) were true but the impact of inequality was estimated based on the miss-specified equation $y_{it} = \delta_1 D_{it} + \eta_i + w_{it}$, with $w_{it} \equiv v_{it} + \delta_2 D_{it-1}$. In case we fit a regression line like that in Figure 2*a*. (i.e., OLS based on differences), the estimated coefficient converges to $\delta_1 - \delta_2(1-\pi)$ as the number of observations goes to infinity; if we do a regression similar to that in Figure 2*b*. (i.e., OLS based on levels), the estimator of δ_1 converges to $\delta_1 + \delta_2 \left(\pi(L^2+H^2) + (1-\pi)(LH+LH)\right)/(L^2+H^2)$. Note that these limits become arbitrarily close to δ_1 and $\delta_1 + \delta_2$, respectively, as π approaches 1. Thus, as already informally argued above, the estimated coefficient reflects the positive short-run relationship, δ_1, in case we rely on first differences while the level-based estimation tends to reflect the negative overall effect, $\delta_1 + \delta_2$.

4 Conclusions

The main point of the present paper is to argue that the empirical literature on the inequality-growth relationship has thus far neglected a strong theoretical implication, namely that the positive and the negative effects of inequality cluster in a very specific way: In our reading, the theoretical literature suggests that the growth-promoting effects tend to set in quickly (i.e., in the short or medium run) while the adverse consequences of inequality require more time to materialize. This observation is important in at least two different dimensions. *First*, with this specific time pattern in mind, we can interpret the existing – and seemingly conflicting – empirical results in a natural way: The differences-based estimation methods (i.e., the FE or first-difference GMM approaches) are prone to systematically pick up the beneficial short- or medium-run implications – and thus tend to indicate a positive relationship. The level-based methods, on the other hand, also reflect the slowly materializing (but powerful) adverse consequences of inequality; thus, the mostly negative results associated with RE or system GMM estimators should be interpreted as the overall effect of inequality in the long run. *Second*, the observation that the positive and the negative consequences of inequality manifest themselves at different points of time has implications for future empirical research: Regression equations including just one (linear) inequality term are likely to be miss-specified. According to our model, an appropriate equation should include several Gini coefficients which control for inequality at different points in the past. Clearly, the successful estimation of such equations requires long time series – and thus may become feasible only in the future.

References

[1] Acemoglu, Daron, Simon Johnson, and James Robinson (2001); "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91(5), 1369-1401.

[2] Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni (2008); "Emergence and Persistence of Inefficient States," mimeo, MIT.

[3] Alesina, Alberto and Dani Rodrik (1994); "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109(2), 465-90.

[4] Arellano, Manuel and Stephen Bond (1991); "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58(2), 277-297.

[5] Arellano, Manuel and Olympia Bover (1995); "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68(1), 29-51.

[6] Banerjee, Abhijit and Esther Duflo (2003); "Inequality and Growth: What Can the Data Say?," *Journal of Economic Growth*, 8(3), 267-299.

[7] Barro, Robert (2000); "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth*, 5(1), 5-32.

[8] Barro, Robert and Jong-Wha Lee (2000); "International Data on Educational Attainment: Updates and Implications," NBER Working Paper No. 7911.

[9] Bénabou, Roland (1996); "Inequality and Growth," in B. Bernanke and J. Rotemberg, eds., *NBER Macroeconomics Annual 1996*, Cambridge, MA: MIT Press.

[10] Benhabib, Jess and Mark Spiegel (1997); "Cross-Country Growth Regressions," mimeo, New York University.

[11] Blundell, Richard and Stephen Bond (1998); "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1), 115-143.

[12] Bond, Stephen, Anke Hoeffler, and Jonathan Temple (2001); "GMM Estimation of Empirical Growth Models," CEPR Discussion Papers 3048.

[13] Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort (1996); "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1(3), 363-389.

[14] Deininger, Klaus and Lyn Squire (1996); "A New Data Set Measuring Income Inequality," *World Bank Economic Review*, 10(3), 565-591.

[15] Sokoloff, Kenneth and Stanley Engerman (2000); "History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 14(3), 217-232.

[16] Foellmi, Reto and Josef Zweimueller (2006); "Income Distribution and Demand-Induced Innovations," *Review of Economic Studies*, 73(4), 941-960.

[17] Foellmi, Reto and Manuel Oechslin (2008); "Why Progressive Redistribution Can Hurt the Poor," *Journal of Public Economics*, 92(3-4), 738-747.

[18] Forbes, Kristin (2000); "A Reassessment of the Relationship between Inequality and Growth," *American Economic Review*, 90(4), 869-887.

[19] Galor, Oded (2009); "Inequality and Economic Development: An Overview," mimeo, Brown University.

[20] Galor, Oded and Omer Moav (2004); "From Physical to Human Capital Accumulation: Inequality and the Process of Development," *Review of Economic Studies*, 71(4), 1001-1026.

[21] Galor, Oded and Joseph Zeira (1993); "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60(1), 35-52.

[22] Glaeser, Edward, Jose Scheinkman, and Andrei Shleifer (2003); "The injustice of inequality," *Journal of Monetary Economics*, 50(1), 199-222.

[23] Heston, Alan, Robert Summers and Bettina Aten (2006); *Penn World Table Version 6.2.* Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania (electronic resource).

[24] Kaldor, Nicholas (1955); "Alternative Theories of Distribution," *Review of Economic Studies*, 23(2), 83-100.

[25] Kuznets, Simon (1955); "Economic Growth and Income Inequality," *American Economic Review*, 45(1), 1-28.

[26] Li, Hongyi and Heng-fu Zou (1998); "Income Inequality is not Harmful for Growth: Theory and Evidence," *Review of Development Economics*, 2(3), 318-334.

[27] Matsuyama, Kiminori (2000); "Endogenous Inequality," *Review of Economic Studies*, 67(4), 743-759.

[28] Persson, Torsten and Guido Tabellini (1994); "Is Inequality Harmful for Growth? Theory and Evidence," *American Economic Review*, 84(3), 600-21.

[29] Perotti, Roberto (1993); "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60(4), 755-776.

[30] Perotti, Roberto (1996); "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Economic Growth*, 1(2), 149-187.

[31] Rosenzweig, Mark und Hans Binswanger (1993); "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *Economic Journal*, 103(Jan.), 56-78.

[32] UNU-WIDER (2008); *World Income Inequality Database, Version 2.0c* (electronic resource).

[33] Voitchovsky, Sarah (2005); "Does the Profile of Income Inequality Matter for Economic Growth?," *Journal of Economic Growth*, 10(3), 273-296.

[34] World Bank (2006); *World Development Indicators 2006* (electronic resource). Washington, D.C.: The World Bank.

Appendix

Proof of Proposition 1. The first step is to introduce some notation. The value function of a representative member of group $i \in \{P, R\}$ is denoted by $V^i(D_t, G_t)$, where D_t and G_t are the two state variables. Thus, when contemplating the preferred level of the public good tomorrow, the poor individuals (i.e., the decisive agents) have to solve the recursive problem

$$V^P(D_t, G_t) = \max_{G_{t+1} \in \{0,1\}} \left\{ a^P \omega^P(D_t) X(G_t) + Z - G_{t+1} F + \beta E \left\{ V^P(D_{t+1}, G_{t+1}) \right\} \right\}.$$

A solution to this problem is a policy function $G_{t+1} = G^P(D_t, G_t)$ which gives tomorrow's level of the public good as a function of the two state variables.

We now start the proof by establishing condition (7) under the premise that the policy function G^P has indeed the properties described in the proposition. The first claim in Proposition 1 is that – irrespective of the value of G_t – the representative poor individual chooses $G_{t+1} = 1$ if $D_t = L$. In formal terms, this means that

$$\begin{aligned} V^P(L, G_t) &= a^P \omega^P(L) X(G_t) + Z - F + \beta \left(\pi V^P(L, 1) + (1 - \pi) V^P(H, 1) \right) \\ &\geq a^P \omega^P(L) X(G_t) + Z + \beta \left(\pi V^P(L, 0) + (1 - \pi) V^P(H, 0) \right), \end{aligned}$$

whereas the second line in the above expression gives the value in case the decision is in favor of the alternative choice, $G_{t+1} = 0$. Then, rearranging terms yields the simple restriction

$$\pi \left(V^P(L, 1) - V^P(L, 0) \right) + (1 - \pi) \left(V^P(H, 1) - V^P(H, 0) \right) \geq F/\beta, \tag{A-1}$$

which is indeed independent of the current level of the public good, G_t. The second claim in the proposition is that – again irrespective of the value of G_t – the poor opt for $G_{t+1} = 0$ if $D_t = H$. Relying on a similar approach as above, it can be shown that this claim requires

$$\pi \left(V^P(H, 1) - V^P(H, 0) \right) + (1 - \pi) \left(V^P(L, 1) - V^P(L, 0) \right) < F/\beta, \tag{A-2}$$

which is again independent of the current level of the public good, G_t.

To proceed further, we have to find expression for the differences in values, $V^P(D, 1) - V^P(D, 0)$, with $D \in \{L, H\}$, that show up in the conditions (A-2) and (A-1). It can be checked that these value differences are given by

$$V^P(D, 1) - V^P(D, 0) = a^P \omega^P(D) \left[X(1) - X(0) \right], \tag{A-3}$$

whereas $D \in \{L, H\}$. Combining expressions (A-1), (A-2), and (A-3) leads then to the condition (7) that is stated in the proposition.

Note, finally, that a similar approach can be used to establish that condition (7) is not only a necessary but also a sufficient condition for the fluctuations equilibrium to exist.