Massachusetts Institute of Technology
Department of Economics
Working Paper Series

INEQUALITY AND GROWTH: WHAT CAN THE DATA SAY?

Abhijit Banerjee, MIT
Esther Duflo, MIT

Working Paper 00-09
June 2000

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

Massachusetts Institute of Technology
Department of Economics
Working Paper Series

INEQUALITY AND GROWTH: WHAT CAN THE DATA SAY?

Abhijit Banerjee, MIT
Esther Duflo, MIT

Working Paper 00-09
June 2000

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

Inequality and Growth: What Can the Data Say?

Abhijit V. Banerjee and Esther Duflo*

June 2000

Abstract

This paper describes the correlations between inequality and the growth rates in cross-country data. Using non-parametric methods, we show that the growth rate is an inverted U-shaped function of net changes in inequality: Changes in inequality (in any direction) are associated with reduced growth in the next period. The estimated relationship is robust to variations in control variables and estimation methods. This inverted U-curve is consistent with a simple political economy model, although, as we point out, efforts to interpret this model causally run into difficult identification problems. We show that this non-linearity is sufficient to explain why previous estimates of the relationship between the level of inequality and growth are so different from one another.

*Department of Economics, MIT, and Department of Economics, MIT and NBER, respectively. We thank Kristin Forbes and Robert Barro for sharing their data, and Alberto Alesina, Oriana Bandiera, Robert Barro, Roland Benabou, Olivier Blanchard, Michael Kremer, Debraj Ray, and Emmanuel Saez for useful conversations.

1 Introduction

It is often that the most basic questions in economics turn out to be the hardest to answer and the most provocative answers end up being the bravest and the most suspect. Thus it is with the empirical literature on the effect of inequality on growth. Many have felt compelled to try to say something about this very important question, braving the lack of reliable data and the obvious problems with identification: Benabou (1999) lists 12 studies on this issue over the previous decade, based on cross-sectional ordinary least squares (OLS) analyzes of cross-country data.

More recently, the literature received a substantial boost from the important work of Deininger and Squire (1996) who put together a much larger and more comprehensive cross-country data set on inequality than was hitherto available. Most importantly, their data set has a panel structure with several consecutive measures of income inequality for each country. This has made it possible to use somewhat more advanced techniques to investigate the effect of inequality on growth: Benhabib and Spiegel (1998), Forbes (1998), and Li and Zou (1998) all look at this relationship using fixed effects estimates, arguing that there are omitted country specific effects that bias the OLS estimates. In contrast, Barro (1999) uses a three-stage least squares (3SLS) estimator which treat the country specific error terms as random, arguing that the differencing implicit in running fixed effects (or fixed effect-like) regressions exacerbates the biases due to measurement errors.

Somewhat surprisingly, both approaches yield new results. While the OLS regressions using one cross-section typically found a negative relationship between inequality and subsequent growth, the fixed effect approach yields a positive relationship between changes in inequality and changes in the growth rate, which has been interpreted as saying that as long as one looks within the same country, increases in inequality promote growth.[1] Barro, by contrast, finds no relationship between inequality and growth. However, he then breaks up his sample into poor and rich countries and finds a negative relationship between inequality and growth in the sample of poor countries and a positive relationship in the sample of rich countries.

To complicate matters further, it is not obvious that these results are comparable. For one, they are based on different data sets: There are relatively few countries that have what

[1] The authors note that this is not necessarily inconsistent with the cross-sectional relationship.

Deininger and Squire (1996) describe as good quality inequality data over a long enough period to make fixed effects estimation possible, whereas the OLS regressions cover more or less the entire world. Moreover, the countries used for fixed effects estimation tend to be the richer countries, which could be related to Barro's claim that inequality has a positive effect in rich countries. Barro himself uses a data set which includes the countries typically used for fixed effects analysis but adds a sizeable number of poorer countries. All these studies also differ in which control variables they include and exclude with the corresponding dangers of controlling for too much or too little. Also, different people have made different assumptions about time-lags which is something that on purely a priori grounds one would expect to make a difference: The OLS regressions typically look at the effect of inequality at a relatively early date (such as 1960) on the next 25 years or more of growth, whereas the fixed effect studies seek to explain growth over five-year periods. Barro takes an intermediate stance, using lags of ten years.

Is there anything then, apart from the obvious fact of disagreement, that we can take away from this body of evidence? This paper, slightly to our own surprise, offers an affirmative answer to this question. Our main conclusion is that there are a set of correlations in the data that do not seem to depend on the choice of data sets and control variables. In particular we find that changes in inequality (in any direction) are associated with lower future growth rates. We also find a strong negative relationship between changes in inequality and past inequality. Finally, there seems to be a negative relationship between growth rates and inequality lagged one period, among countries where the level of inequality was not very high to start with.

This paper stops well short of giving a firm causal interpretation to the relationships we describe: There are too many obvious identification problems. We do, however, suggest that the evidence is reasonably consistent with a simple political economy model that we present in Section 2. This very simple theory predicts relationships that are far from being linear, and the data strongly supports the case for taking the non-linearity seriously. This is in sharp contrast to the uniformly linear models that have been estimated in the literature.

Indeed, this non-linearity is sufficient to explain why different variants of the basic linear model (OLS, fixed effects, random effects) have usually generated very different conclusions: In many cases, it turns out that the differences arise out of giving different structural interpretations to the same reduced-form evidence.

In the end, our paper is probably best seen as a cautionary tale: Imposing a linear structure where there is no theoretical support for it can lead to serious misinterpretations.

The remainder of this paper proceeds as follows. In Section 2, we discuss the different approaches to modeling the relationship between inequality and growth. In Section 3, we present our empirical results. Section 4 shows that these results help us to understand why different methods of estimating the same relationship led to different results. We conclude in Section 5.

2 The Inequality-Growth Relationship

Our goal in this section is to understand what the underlying theory tells us about the appropriate choice of specifications to be used when describing the data on inequality and growth. There are essentially two classes of arguments in the literature that suggest a causal relation between inequality and growth: political economy arguments, and wealth effect arguments. The wealth effect arguments are standard and therefore we limit ourselves to presenting the basic intuition. The political economy arguments we present are somewhat less traditional and therefore we develop them more formally.

2.1 Political Economy Models

Political economy models, in their simplest version, start with the premise that inequality leads to redistribution and we then argue that redistribution hurts growth.[2] Since our goal is to illustrate what can happen in this class of models, we present a version of the argument that minimizes institutional detail.

2.1.1 A Very Simple Model Based on "Hold-up"

Consider an economy constituted of two classes, A and B, which function as competing political groups. Assume that the economy at any point of time is characterized by a single number g which represents the sharing rule for the economy: Group A gets $g\%$ of the output.

In each period this economy is presented with an opportunity which, if availed of, can lead to growth. These opportunities could be a new technology, a trade agreement, an internal

[2] For versions of this argument see Alesina and Rodrik (1994), Persson and Tabellini (1991), Benhabib and Rustichini (1998). For a contrarian point of view, arguing that neither of the two premises of this argument are true in the data, see Benabou (1996b).

reform, or a major foreign investment. The potential growth generated by the opportunity will be denoted by Δy, which is a random variable that is independent over time and has the distribution $F(\Delta y)$.

The growth opportunity does not, however, automatically translate into growth. Some structural changes need to be implemented in order to benefit from the opportunity, and the political system allows for the possibility that these changes would be blocked by one of the groups. To keep matters simple, assume that in every period once the potential growth rate is known, one of the groups, chosen at random, gets to hold up the rest of the economy. More specifically, assume that this group has the option of either acquiescing immediately to the changes, in which case the changes are made and the full growth opportunity is realized, or demanding a transfer from the other group (i.e., an increase in its share) before the cahnges can be made. The other group, in turn, can agree to make the transfer or refuse. If they agree, the changes are made and growth takes place, but by now a part of the growth opportunity has been lost and the economy only grows by $\alpha\Delta y$ ($\alpha < 1$). If the other group refuses to make the transfer, status quo is maintained and there is no growth.

The assumption that there is some efficiency loss in the process of bargaining (i.e., the fact that $\alpha < 1$) plays an important role in our analysis. Delay may be one reason for the loss: It is plausible that the process of getting all members of the losing group to agree to the transfer would take quite some time. Making a credible demand for a transfer typically takes time and resources—as we know, a group might have to resort to industrial action, street protests, and even civil war in order to establish their claim. On the other side, making a credible transfer may require involving third parties (such as the state) and/or changing the institutional framework,[3] which has potential costs of its own. Finally, there are the standard arguments explaining why transfers tend to be distortionary.[4]

To complete the description of the model we assume that all agents are either short-lived or have short horizons. When they decide whether or not to resist, they ignore the effect it will have on output in future periods.

[3] As in, for example, Acemoglu and Robinson (2000).

[4] It must be kept in mind that the transfer could involve abolishing a distortionary tax. For this reason the rest of the examples suggested above fit our purpose better—in those examples, the fact that there is an efficiency loss is independent of the direction of the transfer.

2.1.2 Analysis and Results

Let us assume, without loss of generality, that in a given period it is group B that has the chance to hold up the rest of the economy. Whether or not it does depends on how much it can extract from group A. To figure this out, we need to look at the decision of group A when faced with a demand for transfers worth Δg. If they acquiesce to the transfer their payoff will be $(g - \Delta g)(1 + \alpha\Delta y)$ (the growth rate is $\alpha\Delta y$ because group B has already demanded a transfer). If they do not acquiesce, their payoff will be g, as there will be no growth. Comparing the two, it is clear that the maximum transfer that can be extracted from group A is given by

$$\alpha g\Delta y = \Delta g(1 + \alpha\Delta y).$$

Group B makes its decision taking this as given—it never pays for them to demand more since group A will never acquiesce and there will be less growth in the bargain. They will demand a transfer of size Δg if and only if

$$(1 - g + \Delta g)(1 + \alpha\Delta y) \geq (1 - g)(1 + \Delta y)$$

which implies

$$(1 - g)\alpha\Delta y + \Delta g(1 + \alpha\Delta y) \geq (1 - g)\Delta y.$$

Using the expression for Δg from above this reduces to

$$\alpha \geq 1 - g.$$

Then, $\alpha \geq 1 - g$ is the condition under which group B always demand a transfer when it gets a chance. By a similar argument, the corresponding condition for group A is

$$\alpha \geq g.$$

These two conditions ought to be intuitive: They say that each group will hold up the rest of economy when its share of output is low, which is when they have the least stake in the growth of the overall economy. This is essentially the same reason why the poor in the standard political economy models choose high levels of redistribution even though it hurts growth.

Note also that both of these conditions make no mention of Δy. The potential growth rate for the economy does not influence the probability of growth-reducing bargaining/conflict. The growth rate in our economy only depends on whether there is a hold-up: If there is no hold-up the rate is Δy, while if there is a hold-up it is $\alpha\Delta y$. In the world of this model, hold-ups only happen when there are redistributive transfers that result from the hold-up. Therefore:

Result 1: The growth rate in this economy in any period following a distributional conflict (i.e., hold-up) is lower than when there is no conflict.

The data we will use does not give us direct measures of hold-up. Observe, however, that in our model there is a perfect correspondence between hold-ups and distributional changes at the onset of the growth episode, and we do have measures of those distributional changes that show up as changes in measured inequality. We therefore want to interpret the variable g as a measure of inequality. This is possible if we are prepared to assume that one of the groups (say group A) is substantially richer than the other in terms of per capita income (in other words, group B has a much larger share of the population than group A). In this case, an increase in g in our model would correspond to an increase in inequality.[5]

The relationship between distributional changes and growth implied by the above result is, however, highly discontinuous. This is because our model clearly makes an excessively strong distinction between the case where there are no distributional changes and the case where there are some distributional changes. A smoother relationship could be derived if we assumed instead that the hold-up problem only determines the planned transfer, whereas the actual transfer is determined ex post by random forces. Combined with the assumption that growth is higher when there is less actual transfer, this would give us *an inverted U-shaped relation between growth and distributional changes.* Growth is maximized when there are no changes and is lower when there are changes in either direction.

If we were prepared to take this model literally, it would allow us to estimate a (non-linear) causal relationship between growth and changes in inequality. There are, however, many reasons why this model is special: Most importantly perhaps, growth here does not have any direct distributional effect. If more growth leads to more redistribution, then the anticipation of a large growth shock could raise the likelihood that there is a hold-up problem. More redistribution could then be associated with higher growth and the relationship would no longer be U-shaped. More importantly, there would be reverse causality—running from growth to anticipatory changes in the distribution—making it impossible to interpret the relationship between

[5]This interpretation clearly only makes sense if g is not too small.

growth and distributional changes causally.[6] We therefore only offer this model as a possible way to interpret the data.

The discussion above suggests that, at least in terms of data description, if not causal interpretation, we should estimate a relationship of the form:

$$\frac{(y_{it+a} - y_{it})}{a} = \alpha y_{it} + X_{it}\beta + k(g_{it} - g_{it-a}) + v_i + \epsilon_{it} \ , \qquad (1)$$

where y_{it} represents the logarithm of GDP in country i at date t, a is the length of the time period we choose, 5 or 10 years in the examples we will consider ($(y_{it+a} - y_{it})/a$ is therefore the growth rate). X_{it} is a set of control variables, g_{it} is the gini coefficient in country i at date t, and $k(\cdot)$ is a generic function. At this point we do not impose any structure on the shape of the $k(\cdot)$ function. The error term is modeled as a country-specific time invariant effect (v_i) and a time varying error term (ϵ_{it}). y_{it} is included among the controls in order to capture convergence effects, and X_{it} to control for possible sources of spurious correlation.

However, the political economy literature has not taken this route. Instead, the approach taken has been to derive a relationship between the level of inequality and changes in inequality, which, combined with a relationship between level of inequality and changes in inequality (such as the one just derived), generates a relation between growth and the level of inequality.[7] We could also take a similar approach here. To do this, observe that in our model changes in inequality are causally related to the level of inequality. If $g > \alpha$, group A will not hold up the economy even if it has the chance. If $g < 1 - \alpha$, group B will similarly desist. Therefore g will increase when it is low and fall when it is high. What happens in the middle depends on whether $\alpha < \frac{1}{2}$. When this condition is satisfied, for values of g around $\frac{1}{2}$, there will be no hold-up and inequality will not change. By contrast, if $\alpha > \frac{1}{2}$, there will be an interval of g values centered around $\frac{1}{2}$, where each side will try to hold up the other and inequality is equally likely to move in either direction. The expected change in inequality is once again zero. We summarize the implications of this discussion in:

[6] Note that we are not worried about the direct effect of growth on distribution (the Kuznets curve effect) because that is presumably subsequent or contemporaneous to the growth episode. What worries us is the fact that there may also be an effect on the distribution *prior* to the growth episode.

[7] See Alesina and Rodrik (1994), Persson and Tabellini (1991) and Alesina and Perotti (1996). The argument in Alesina and Perotti (1996) is most closely related to ours: Income inequality leads to political instability and hence to lower growth; indeed, instability may be a symptom of what we call grabbing.

Result 2: The relation between the level of inequality and the expected change in inequality in our model is broadly negative: It is strictly decreasing at both extremes and flat at intermediate levels of inequality.

This suggests estimating the following relationship:

$$g_{it+a} - g_{it} = \alpha y_{it} + X_{it}\beta + h_1(g_{it-a}) + v_i + \epsilon_{it}. \tag{2}$$

What matters for growth in our model, however, is not the actual change in inequality but the absolute value of that change (as both positive and negative changes reduce growth). The shape of this relationship is also implied in our discussion above: If $\alpha < \frac{1}{2}$, there are no hold-ups at intermediate values of inequality and therefore inequality does not change at all. Growth is therefore maximized. By contrast when $\alpha > \frac{1}{2}$, there will be changes in both directions at intermediate levels of inequality and therefore growth will be slowest in this range. We state these conclusions as:

Result 3: The relation between the level of inequality and absolute changes in inequality for the economy in our model is U-shaped when $\alpha < \frac{1}{2}$, i.e. there are more changes when inequality is either very high or very low. If $\alpha \geq \frac{1}{2}$, the shape implied by our model is inverted U-shaped, i.e., there are more changes at intermediate levels of inequality.

Combining this with Result 1 gives us:

Result 4: The relation between the level of inequality and future growth for the economy in our model is inverted U-shaped when $\alpha < \frac{1}{2}$, i.e., there is less growth when inequality is either very high or very low. If $\alpha \geq \frac{1}{2}$, the shape implied by our model is U-shaped, i.e., there is less growth at intermediate levels of inequality.

Within the world of our model the level of inequality is indeed exogeneous and therefore we could, in principle, estimate the following empirical specifications which correspond to results 3 and 4.

The first relationship relates the square (or, alternatively, the absolute level) of changes in inequality to the level of inequality.

$$(g_{it+a} - g_{it})^2 = \alpha y_{it} + X_{it}\beta + h_2(g_{it-a}) + v_i + \epsilon_{it}. \tag{3}$$

The second relationship is a "reduced form relationship", which relates the level of inequality (lagged one period) to the growth rate:

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + h(g_{it-a}) + v_i + \epsilon_{it}, \tag{4}$$

where once again $h(\cdot)$ may be non-monotonic.

It is worth noting that estimating these relationships using cross-country data introduces a number of additional problems. First, α may be different for different countries and therefore the shape of the relationships may vary across countries: They may be U-shaped in some and the reverse in others. Second, the value of measured inequality that corresponds to $g = \frac{1}{2}$ may vary from country to country, and therefore the relationship may peak (and bottom out) at different points in different countries. For both these reasons the relationship estimated from cross-country data may be very different from the relationship in any single country.

It remains, however, that the correspondence between results 3 and 4 should hold even when these countries are heterogenous. In other words, as long as our basic model is correct, it is always a prediction of our model that our estimates of the functions $h(\cdot)$ and $h_2(\cdot)$ in equations 4 and 3 should be mirror images of each other.

2.1.3 Discussion

The goal of the political economy models that are standard in the literature is to derive and estimate a relationship that corresponds to our equation 4. However, while we have emphasized the non-linearity and non-monotonicity of the predicted relationships, they have typically derived a monotonic relationship that they estimate using a linear model. The difference arises from the fact that we do not make the common assumption that it is only redistributing to the poor that is costly. This assumption is a natural consequence of assuming that the main cost of redistribution comes from the waste that results from high taxation. Our view, by contrast, is that redistributing a significant amount in either direction is almost always costly, since it usually comes with some degree of upheaval.

Our conclusion that changes in inequality are associated with lower growth is based on the immediate effect of these changes. The long-run effect may be very different: Within our model, the long run effect of changes in inequality is to move the economy towards intermediate levels of inequality. When $\alpha < \frac{1}{2}$, intermediate levels of inequality are associated with the highest growth rates: Changes in inequality may therefore be good for growth in the long run. Going beyond our model, it is easy to imagine a redistribution that, for example, takes the form of more investment in public schooling today could, in the long run, raise the growth rate.

Finally, our model is quite special in its prediction that all changes in inequality are bad for growth. One could easily imagine a variant of the model where, for example, there are periods where there are no growth opportunities and no changes in inequality. In such a model, periods where there are small changes in inequality may be associated with more growth than periods where there are no changes whatsoever.

2.2 Wealth Effect Arguments

The basic idea behind the wealth effect arguments is that there is a concave relationship between the current wealth of an individual and his future wealth. Such a concave relationship can be generated by assuming credit-market imperfections: Intuitively, the poor under-invest because credit markets are imperfect and as a result earn higher average returns on their wealth than the rich, who invest more (see for example Benabou (1996a), and Bardhan, Bowles and Gintis (2000)). When the relationship is concave, a mean preserving spread in the distribution of current wealth reduces the mean wealth in the future, implying a slower growth rate. In other words, inequality is negatively related to growth.

There are, however, a number of reasons why this relationship need not always be so well-behaved. As pointed out by Galor and Zeira (1993), non-convexities in the investment technology have the effect of reversing this relationship over some ranges. The reason is easy to see at least in the case where there is a minimum scale for investment: If this minimum were high enough, the average person in an economy would not be able to afford it (because of constraints on how much he can borrow) and as a result there would be no investment and no growth in this economy if everyone owned the mean wealth. Some differentiation is unavoidable in this economy if there is to be growth.

Elaborating this model to allow for the endogenous determination of factor prices (as in Banerjee and Newman (1993), Aghion and Bolton (1997) and Piketty (1997)) opens up further possibilities. In this case the relationship between inequality and growth does not even have to be continuous: Small changes in inequality can lead to large changes in the growth rate.

For all these reasons, we feel that this model is captured best by the following relatively flexible specification:

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + h(g_{it}) + v_i + \epsilon_{it}. \qquad (5)$$

2.3 Nesting the Two Models

Notice that equation 5 is essentially identical to equation 4 with the one difference being that g_{it-a} has been replaced by g_{it}: In other words, both models generate almost the same reduced forms. The reduced forms are exactly the same when, as in Barro (1999), equation 5 is specified using g_{it-a} as the independent variable. This explains why the empirical literature does not often try to distinguish between the alternative models. However, in one case this equation is the basic relationship while in the other it is something that follows from the basic relationship. One could also nest the two models by combining the two basic relationships to get:

$$\frac{(y_{it+a} - y_{it})}{a} = \alpha y_{it} + X_{it}\beta + h(g_{it}) + k(g_{it} - g_{it-a}) + v_i + \epsilon_{it}. \qquad (6)$$

Starting from this general equation, one can impose various restrictions (linearity of$h(\cdot)$ and $k(\cdot)$, leaving out one of these two functions) which would give rise to the models that have been estimated in the literature.

Note that g_{it-a} is strongly correlated with g_{it}. Therefore, equation 4 is a potential reduced form of this general model.

3 Estimation and Results

In the next two sections, we discuss the estimation of equations 4 and 6. We start by describing the data using flexible estimation methods. We then discuss the consequences of our findings on the interpretation of the results in the literature.

Our main focus in this paper is the potentially non-linear effects of distributional changes and therefore we have chosen to sidestep a number of important and natural questions. First, we do not choose a new set of control variables. The choice of these variables is clearly critical, since a central concern for the empirical literature is that the gini coefficient could proxy for omitted variables. For example, Barro (1999) criticizes earlier studies on their choice of control variables and shows, in particular, that their results are sensitive to the inclusion of fertility in the regression. But the choice of the variables entails making judgements about causality that are not easy to defend. We therefore avoid taking a position on this subject. Instead, we present all the results for the set of control variables (X_{it}) used in Perotti (1996) and the set of control variables used in Barro (1999). These specifications are useful benchmarks for two reasons. First, the Perotti specification has been used by most subsequent studies. Second, they represent two extremes: The Perotti specification uses the smallest number of control variables and the Barro specification the largest. The list of variables included in both specifications is included as a note to Table 2. The Perotti specification excludes most variables (in particular, investment and government spending) through which the influence of inequality could be channeled. The only variables included are male and female education and the purchasing power parity of investment goods, a measure of distortions. Barro, on the other hand, includes investment share of GDP, fertility, education, and government spending, which are plausible channels through which inequality could affect growth.[8] The interpretation of the coefficient of inequality in the two regressions is therefore different.

Second, we do not experiment with alternative definitions of inequality (interquartile range, measure of poverty, etc.). There are reasons to doubt that the gini coefficient is the appropriate measure of inequality. However, most empirical work on growth and inequality focuses on the gini coefficient. Therefore, our focus in this paper is also on the relationship between the gini coefficient and economic growth. A distinct but related question concerns the reliability of the measure of the gini coefficient. A new data set, compiled by Deininger and Squire (Deininger and Squire (1996)), has substantially improved the reliability and the comparability of available measures of inequality. They have compiled an extensive data set for a large panel of countries. They also identify a sub-set of their data as a "high quality" data set.[9] Most recent studies have

[8]In addition, Barro includes the average growth of terms of trade over the period, indices of democracy and the rule of law, the square of the logarithm of GDP, the square of the democracy index, and the average inflation in the period. He implements a three stage least squares method, where he uses lagged values of the regressor as an instrument for current values. As inequality is an instrument for itself in his specification, we will focus on the reduced form and use the instruments as control variables. In particular, we follow Barro and control for y_{it-a}, not y_{it}, in the regression (although this does not affect our results to control for y_{it} instead).

[9]The high quality data set includes only those observations which satisfy the following criteria: The survey comes from a national coverage, the information is based on direct surveys of incomes, the surveys sample the

used this new high quality data set. Therefore, we will present most results in the Deininger and Squire high quality data set restricted to countries with at least two consecutive observations.[10] Using this data set should ensure relatively reliable and comparable measures of inequality, although the data has not been exempt from criticism.[11] It should be noted that, depending on the data source, the data refers either to ex post inequality (i.e., to income measured net of redistribution, or to expenditure inequality) or to gross inequality. The distinction is less strong than it appears, however, since a substantial fraction of the redistribution does not occur through the tax system but through other mechanisms (minimum wages, labor laws, inflation, etc.).

An additional drawback is that the high quality data set is small, and includes very few poor countries, especially when it is limited to countries where at least two observations are available. In an attempt to expand the sample size, Barro proposed adding some observations that were rejected by Deininger and Squire on the grounds that they were not identified by a clear primary source. The coverage increases substantially, at the expense of an additional reduction in the accuracy of measurement. We will also present some results using the Barro sample for comparison with his results.[12] In Table 1, we present selected descriptive statistics for both samples. Countries in the Barro sample tend to be poorer and to have higher levels of inequality. A number of them are located in sub-Saharan Africa. Appendix A contains a list of the countries in each sample.

Third, the question of the relevant time period (the choice of a) is also important. As

complete population (not only those earning an income), the data does not come from tax records, and, finally, the data gives a clear reference to the primary source.

[10]This is the sample used in Forbes (1998) and Li and Zou (1998). The Deiniger and Squire data set provides the year in which the observation was taken. To construct a measure of inequality every 5 years, we follow Forbes (1998) and we chose the closest measure in the 5 years preceding the relevant date if the measure was not available for this particular year. We also follow previous studies in adding 6.6 to the gini when it was constructed from expenditure instead of income. However, still following the other studies, we did not attempt to correct the gini coefficient for whether it was gross or net of taxes, and whether the unit of measurement was the household or the individual.

[11]See Atkinson and Brandolini (1999). In particular, they argue that, for the OECD countries at least, synthetic or tax-based measures of inequality, rejected from the high quality data set, would permit the construction of more consistent inequality series than the estimates included in Deininger and Squire's data set.

[12]For the Barro sample, we use the data used by Barro. Barro constructed inequality measures for 1960, 1970, and 1980 from the D-S data set, so we use the Barro sample only in 10-year periods regression.

we emphasized in the previous section, the theory predicts different effects over different lags. The first set of empirical papers studied the growth over a long time period (25 to 30 years). Subsequent papers have exploited the richness of the Deininger and Squire data set and have chosen shorter lags (5 or 10 years) in an attempt to increase the number of available observations. Since using longer lags substantially reduces the number of changes in inequality in our data set, we will focus on 5 year lag periods. Using the expanded Barro data set, we will also explore the sensitivity of our results when choosing 10 year lags.

3.1 Basic Results

Table 2 presents the results from estimating various versions of equation 6.[13] In columns (1), (2), (7), and (8) we suppress the $k(\cdot)$ function, and regress the growth rate on the gini coefficient. As in previous work (e.g., Forbes (1998)), we do not find any effect of the gini coefficient in this specification. The relationship does not seem to be particularly non-linear: A squared term introduced in the regression is not significant either, and the two terms are not jointly significant.

In columns (3) and (9), we instead suppress the $h(\cdot)$ function, and we regress growth on the change in inequality and the change in inequality squared. By contrast to the previous results, past variation in inequality is related to subsequent growth, in a very non-linear way: While the linear term is insignificant, the quadratic term is negative and significant with both sets of control variables.

We then introduce the level of the gini coefficient into the regression (columns (4) and (10)). The coefficients of $(g_{it} - g_{it-a})$ and $(g_{it} - g_{it-a})^2$ are not affected by the introduction of the gini coefficient.[14] To explore the non-linearity further, we use a kernel regression, and we "partial out" the linear part of the model (i.e., y_{it}, g_{it} and X_{it}) using a method analogous to that developed by Robinson (1988) and applied in Hausman and Newey (1995).[15] The results

[13]All of these equations are estimated using a random effect specification, to allow for correlation of growth rates between countries over time.

[14]We present the results with only a linear term in the gini coefficient because we did not find any strong non-linearity when we looked at the $h(.)$ function separately, but the exact same results are obtained if we introduce higher-ordered polynomials as well.

[15]This is implemented by first regressing all the control variables (y_{it}, g_{it} and X_{it}) and the dependent variable $\Delta y_{it+a} = y_{it+a} - y_{it}$ non-parametrically on $\Delta g_{it} = g_{it} - g_{it-a}$ and forming the residuals of this non-parametric regression. Estimates of the parameters α and β are then obtained from the OLS regression of the residual of

are shown in Figures 1A (with Perotti variables) and 1B (with Barro variables). The kernel regression line is shown as a solid line. This relationship has the shape of an inverted U, with a maximum around 0 and a relatively flat section at the top. Changes in inequality, in any direction, are associated with reduced growth in inequality, and larger changes are associated with larger decline in growth.

This result is striking, and we investigated its significance using a variety of methods. First, we estimated the relationship using series estimation. In Figure 1, we show the predicted value using a quartic specification for the function $h(\cdot)$. This polynomial is maximized when the value of lagged change in inequality is 0.012 (using Perotti variables), which is very close to 0. To test whether the non-linearity is statistically significant, we present in columns (6) and (12) the F test for the joint significance of the non-linear terms in the partially linear model. Linearity is rejected in both cases, at 3% in the Perotti specification and 12% in the Barro specification. Given the limited amount of data (128 and 98 observations, respectively) and the fact that it is very noisy, this result is a surprisingly strong rejection of linearity. Finally, we estimate a piece-wise linear specification for $h(\cdot)$ (columns (5) and (11)), where we treat the effects of increases and decreases in inequality separately. The coefficients of decreases and increases in inequality are positive and negative, respectively. The positive coefficient in the decreasing range is significant in both specifications. The negative coefficient in the increasing range is significant only in Perotti's specification. We also ran these specifications using the Barro expanded data set, and 10 year lags instead of 5 year lags, and we find the same inverted U-shaped relationship between changes in inequality and growth, albeit estimated with less precision, whichs is not surprising given that we are left with only 78 observations.

On balance, there is no strong evidence of a direct correlation of inequality on growth in the short run (over a 5 year lag period), but there seems to be an association between changes in inequality and growth. Changes in inequality, *whatever their direction*, are associated with lower growth in the next period. We discuss at the end of this section whether any causal interpretation can be given to this result, but before that we report the results from our reduced form estimates.

---

the dependent variables on the residuals of the control variables. Finally, the function $ah(.)$ is estimated by estimating non-parametrically the function: $\hat{E}(\Delta y_{+a}|\Delta g)$, $\hat{E}(\Delta y|\Delta g)$, $\hat{E}(X|\Delta g)\hat{\beta}$ and forming the difference $\hat{E}(\Delta y + a|\Delta g) - (a\hat{\alpha} + 1)\hat{E}(\Delta y|\Delta g) - a\hat{E}(X|\Delta g)\hat{\beta}$.

3.2 Reduced Form Results

In Table 3, we present the results of the estimation of equation 4. The difference between the specifications estimated in this table and the first columns in the previous table is that the independent variable is not the beginning-of-period level of inequality ($g(t)$) but the *lagged* level of inequality ($g(t-5)$).

The coefficient of $g(t-a)$ entered linearly is now negative (around -4%), but still insignificant in both Perotti's and Barro's specification (Table 3, columns (1) and (4)). The partially linear model (shown in Figure 2 for Perotti's control variables) seems to indicate a non-linearity: The derivative seems to switch signs from negative to positive around a value of the gini coefficient ranging between 0.40 and 0.45. However, we cannot statistically reject that a linear specification with a coefficient of 0 characterizes the data (columns (2) and (5)). Columns (3) and (6) show the results obtained when we lag the other regressors by one period as well, which, as we show below, is similar to the the reduced form of the models of Barro (1999) and Forbes (1998). The coefficient of lagged inequality is similar in these specifications. It is significant with the Barro control variables.

It is interesting to compare these results with those of Barro (1999), who estimates a very similar equation. He wants to explain the growth over the three decades (1965-1975, 1975-1985, 1985-1995) for a sample of countries. As his independent variable, he uses inequality lagged by 5 years. For example, for the decade 1965-1975, inequality *in or around 1960* (as opposed to 1965) is used as an explanatory variable. He estimates a structural equation of the form:

$$(y_{it+2a} - y_{it})/2a = \alpha y_{it} + (Z_{it+2a})\beta + \delta g_{it-a} + \nu_i + \xi_{it} \ , \tag{7}$$

where Z_{it+2a} is a set of control variables (mostly period averages). The instruments used for Z_{it+2a} in equation 7 are a set of variables fixed over time and the lagged values of most regressors. Barro's reduced form can therefore be written as:

$$(y_{it+2a} - y_{it})/2a = \lambda y_{it-a} + X_{it}\kappa + \delta g_{it-a} + \nu_i + \xi_{it} \ , \tag{8}$$

which is very similar to equation 4.[16] The only differences are that the growth is averaged over a decade and that Barro used a larger sample.

[16] Note that we have used y_{it-a} instead of y_{it} in our estimates of Barro's specification.

Columns (1) and (2) of Table 4 report the estimates of equation 8, in Barro's sample using 10 year lag periods for Barro's and Perotti's control variables, respectively. These coefficients are similar in magnitude to what we had previously estimated based on equation 3: Negative and not significant point estimates of -0.026 using Perotti's variables and and -0.023 using Barro's variables.

Barro's reduced form and the structural form coefficients have a relatively direct relationship, as g_{it-a} serves as an instrument for itself.[17] However, his point estimate of the structural coefficient is close to 0. In Table 4 (columns (3) and (5)), we report our replication of Barro's structural estimate, with and without controlling for fertility. As in Barro, we find a coefficient of 0 when fertility is included, and the results are similar to other results in this paper when it is excluded. The non-parametric regression without fertility is shown in Figure 3. The relationship is even more non-linear than with the Perotti variables.

The larger sample used by Barro makes it possible to investigate the issue of non-linearity further by breaking the sample. It turns out that most countries on the right of the turning point (gini above 0.45) in the rich country sample are in Latin America. Therefore, we examine the difference between the coefficient of the gini coefficient both in and outside Latin America. The results are presented in columns (4) (with fertility) and (6) (without fertility). In columns (7) and (8) we present the corresponding reduced forms. High levels of inequality are associated with higher subsequent growth in Latin America, while they are associated with lower subsequent growth in the rest of the sample. In the reduced forms, both the negative coefficients outside Latin America and the positive coefficients within Latin America are significant. The magnitude of the coefficient outside Latin America is slightly larger than the coefficients we obtained in Table 3. The coefficients are not affected by the inclusion of fertility, or fertility interacted with a dummy for Latin America in the set of control variables. The non-linearity seems to be mainly due to the difference between Latin America and the rest of the sample. The relationship is fairly linear (but of opposite sign) in both sub-samples (Figures 4A and 4B). Looking separately at rich and poor countries, outside Latin America, yields similar results.

We conclude that inequality, lagged one period, is negatively correlated with growth in countries in the range where it is not too large (below 0.40), and, in particular, outside Latin America.

[17]The only difference between the reduced form and the structural form coefficient of inequality is that inequality also serves to instrument for other variables in equation 7.

Note that results are sensitive to the lag structure: We found no relationship whatsoever between base-period inequality and growth. We come back to the interpretation of this difference in section 3.4.

3.3 Results about Changes in Inequality

In columns (7) to (10) of Table 3, we estimate the relationship between changes in inequality and past inequality described by equations 2 and 3. In both the Perotti and Barro specifications, changes in inequality is strongly negatively correlated with past inequality, while the square of the change in inequality is positively related to inequality.

The corresponding kernel regressions are shown in Figures 5A and 5B. The relationship between inequality and changes in inequality is fairly linear. In contrast, as Figure 5B clearly demonstrates, the relationship between inequality and squared changes in inequality is non-linear with a peak around 0.45. The shape is very similar if we replace the square of the change with its absolute value.

3.4 Interpretation

The wealth effect model had suggested looking at the correlation between growth and the level of inequality. Our results in Table 2 show the effect of contemporaneous inequality on growth is never significant. One might argue, however, that we should look at the effect of lagged inequality instead since, for example, the wealth effect might work by raising investment in human capital which affects growth with a lag. The specification in Table 3 replaces the base-period with the inequality lagged 5 years, and the results are only a little bit more promising: The coefficient is always negative but almost never significant.

The political economy model had suggested looking at the correlation between growth and changes in inequality: This relationship is indeed U-shaped, as predicted by the model (see result 1 above). The relationship between changes in inequality and the level of inequality, shown in Figure 5A is clearly negative, which is consistent with our result 2.

Figures 2 and 5B, taken together, are also consistent with the prediction (of our model) that they should be mirror images of each other: Figure 2 shows the relationship between growth and the lagged level of inequality. The relationship seems to be U-shaped: High inequality is associated with lower growth, but, once the gini coefficient is higher than 0.40, further increases are associated with increases in the growth rate. Figure 5B shows the relationship between the square of changes in inequality[18] and lagged inequality. This relationship is non-linear as well: Higher levels of inequality are associated with higher changes, but the relationship turns around when the gini coefficient reaches 0.45. There is a surprisingly close match between the shape of the curve in Figure 5B and the curve shown in Figure 2. In particular, both curves have more or less the same turning point (between 0.4 and 0.45). This suggests the possibility that the reduced form relationship might be driven, at least in part, by the relationship between current inequality and future changes in inequality.

Our model does not have very strong prediction on the shape of each of these two relationships, taken one at a time. Even if all countries were identical in all respects, depending on the value of α, Figure 2 could be either U-shaped or the inverse. The fact that we find a U-shape may indicate that α is bigger than $\frac{1}{2}$ (redistribution is not too costly) but it could also mean that the countries are very different.

3.5 Identification Problems

We do not, however, want to claim our evidence clinches the case for the political economy model: It is an interpretation of the reduced form which cannot be rejected by this data. We now turn to other possible interpretations of this evidence.

First, the data is measured with a considerable amount of error. The idea that measurement errors can explain some "findings" in cross-country growth regressions has been put forward by Krueger and Lindahl (1999) in the case of education, and Barro (1999) in the case of inequality and growth. Therefore, it is worth examining whether it could explain the patterns we have seen in the data. The relationship between inequality and changes in inequality (Figure 5A) could be due to mean-reversion induced by classical measurement error. However, classical measurement error should not lead to the inverted U-shaped relationship between changes in inequality and subsequent growth. But, except for convenience, there is no reason to assume that measurement errors are always of the classical sort. Imagine that true inequality does not change quickly. Statistical agencies add noise to the true measure of inequality.[19] The more

[18]The absolute value would gives the same result.

[19]The Deininger and Squire series are not all fully consistent over time. For example, their data for France

competent the agency, the less noise there is in the data. The assumption seems plausible, and immediately delivers the conclusion that in countries with more competent statistical agencies, large changes in inequality will be less frequent than in countries with less competent statistical agencies. Assuming that, all else being equal, countries with incompetent statistical agencies tend to grow more slowly, we can explain the non-linear effect of changes in inequality on growth in a specification which does not control for past growth.

Below, we will discuss a different specification (the results are presented in Table 5) where we control for lagged value of growth, and therefore, for any characteristics of the country that are fixed over time. We find the inverted U-curve even with this specification. To attribute the result entirely to measurement error, we would then need to assume that changes in the competence of the statistical agency are correlated with innovation in growth, which is less palatable. Further, a direct examination of the list of countries with high absolute changes in inequality, which we present in Table A2 in the appendix, indicates that some countries with fairly well developed statistical systems experienced important changes in inequality. We are, therefore, reluctant to commit ourselves to even such an "enriched" measurement error explanation of the inverted U-shaped relationship between changes in inequality and growth, but we acknowledge it as a possible interpretation.

Second, even if the inequality is properly measured, there are several reasons why the non-parametric relationships between inequality or changes in inequality and growth may not be causal. In discussing variants of our basic model, we had already suggested a possible source of reverse causality. An alternative source of reverse causality comes from the idea that the lack of growth opportunities makes the environment more conflictual (say, because people feel frustrated), and conflicts lead to changes in inequality.

More generally, neither changes in inequality nor levels of inequality are randomly assigned and one can easily imagine reasons why they may be spuriously correlated with growth. For example, greater social and political instability may lead to both lower growth and volatile levels of inequality. During episodes of growth due to fast technological change, the distribution may also greatly change.[20]

shows a sharp drop in inequality from 1975-1980. As Atkinson and Brandolini (1999) show, this is due to a rupture in the series rather than to a genuine change in the underlying inequality.

[20]This, of course, gives us a pattern opposite to the one we find, with high growth correlating with large changes

Going beyond the political economy models, one could argue that cultural structures (such as a caste system) may restrict occupational choices and therefore may not allow individuals to make proper use of their talents, causing both higher inequality and lower growth. Alternatively, it may be because countries that (for exogenous reasons) use technologies intensive in skilled labor, have both more inequality and faster growth. Countries also differ in their financial institutions. The standard wealth effect arguments would then predict that the country with the better capital market is more likely to be more equal and to grow faster (at least once we control for the mean level of income). The correlation between inequality and growth will, therefore, be a downwards biased estimate of the causal parameter. In addition, if the causal effects of inequality vary across countries (because, for example, it is related to the level of financial development, as suggested by Barro (1999)), the OLS coefficient is a weighted average of different parameters, where the weights are the country-specific contributions to the overall variance in inequality. It is not at all clear that we are particularly interested in this specific weighted average.[21]

Finally, the fact that we argue the results are broadly consistent with the political economy model does not rule out the possibility that they are entirely driven by wealth effects. The mean reverting pattern between inequality and changes in inequality (Figure 5A), for example, would be generated by any model where the long run distribution is independent of initial conditions (as in Loury (1981), for example). As we have already noted, the wealth effect models can also generate a highly non-linear relationship between inequality and growth. The implied relationship between changes in inequality and growth may then also be non-linear, and could turn out to have exactly the U-shape that we have found.

4 Relationship with the Literature

Regardless of the interpretation of these features of the data, it is clear that they have important implications for how we interpret the variety of results in the literature. In particular, we will show that the striking results obtained by those who have estimated the growth-inequality

in inequality. However, it remains that the estimated relationship is biased.

[21]The difference can be important: Deaton and Paxson (1998) show that the estimated relationship between savings rates and the dependency rates changes dramatically (from negative to positive) when each country's specific coefficient is weighted by the GNP of the country instead of the OLS weight.

relationship with fixed effects arise from giving a different and misleading interpretation to the same reduced form evidence that is presented here.

4.1 Parametric Approaches in the Literature

The standard procedure for estimating the relationship between inequality and growth in the literature is to ignore $k(g_t - g_{t-a})$, and to assume that $h(g) = \gamma g$, which gives us

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + \gamma g_{it} + v_i + \epsilon_{it}. \tag{9}$$

If this is indeed the real structure of the data, it is possible to solve some of the identification problems raised in the previous section. Essentially, taking out period averages of variables eliminates the (additive) country fixed effect, thus allowing the interpretation of the fixed effect coefficients as the causal effect of inequality on growth under the assumption that the innovation in the error term is not correlated with changes in inequality.

Alternatively, one could first difference equation 9:

$$\frac{(y_{it+a} - y_{it})}{a} - \frac{(y_{it} - y_{it-a})}{a} = \alpha(y_{it} - y_{it-a}) + (X_{it} - X_{it-a})\beta + \gamma(g_{it} - g_{it-a}) + \epsilon_{it} - \epsilon_{it-a}. \tag{10}$$

This is a relationship between changes in the gini coefficient and changes in the growth rate. As long as $\alpha = 0$, the OLS estimate of this relationship gives an unbiased measure of α and is statistically equivalent to the fixed effect estimate of equation 9.

One problem is that when α is not equal to zero, the presence of lagged dependent variables on the right-hand side biases the OLS estimate of the differenced equation (as well as the fixed effect estimate of equation 9). The literature (notably Forbes (1998) and Benhabib and Spiegel (1998)) has then followed the lead of Caselli, Esquivel and Lefort (1996) in using a GMM estimator developed by Arellano and Bond (Arellano and Bond (1991)). The idea is to multiply equation 9 by a, to put y_t on the right side, and to take first differences of the resulting equation. This leads to the following equation:

$$y_{it+a} - y_{it} = (a\alpha + 1)(y_{it} - y_{it-a}) + a(X_{it} - X_{it-a})\beta + a\gamma(g_{it} - g_{it-a}) + a\epsilon_{it} - a\epsilon_{it-a}. \tag{11}$$

An unbiased estimate of γ can be generated if this equation is estimated using y_{it-a}, X_{it-a}, g_{it-a} and all earlier lags available as instruments for $(y_{it} - y_{it-a})$, $(X_{it} - X_{it-a})$ and $(g_{it} - g_{it-a})$.

Results of estimating equation 9 by random effects, fixed effects, first difference, and Arellano and Bond estimators are presented in Table 5, for 5 year periods in the Deininger and Squire high quality sample. The results are very consistent. First differences, fixed effects, and Arellano and Bond coefficients are positive and significant in both specifications. Forbes (1998) and Li and Zou (1998), who first made this observation, have shown that this result is robust to a wide variety of changes in specifications.[22] Li and Zou (1998) propose a theoretical explanation based on a political economy model. Forbes (1998) rightly notes that the estimated coefficient indicates a short-run positive relationship between growth and inequality, which might not directly contradict the long-run negative relationship. She points out that her results suggest that "in the short and medium term, an increase in a country's level of income inequality has a significant and positive relationship with subsequent economic growth".

Taking out fixed effects exacerbates the measurement error problem, especially for a variable like the gini coefficient, for which the variation across countries is more important than the variation over time. Classical measurement errors alone should not, however, explain why the coefficient of inequality should change signs, becoming positive and significant. Furthermore, the GMM estimator instruments first differences with lagged levels, which should, in principle, attenuate the classical measurement error problem. Therefore, there is probably more to this reversal in sign than simple measurement error.

4.2 Non-Linearity

Note that all these approaches rely heavily on the linearity of equation 9 and the exclusion of the differenced term. If either of these conditions are violated, the fixed effect and first difference estimates of γ will not be identical, and both will be different from the OLS estimate of equation 9 even if all the other conditions for the validity of the OLS estimate are satisfied. It will then be important to be very careful in interpreting each of these coefficients.

The results in the previous section suggest that changes in inequality were negatively correlated with subsequent growth. Assuming the relationship between the level of inequality and

[22] We were not able to exactly replicate Forbes (1998) result for the Arellano and Bond estimator (she obtains with the Perotti specification a coefficient of 0.13). The coefficients of the other regressors are similar.

growth is indeed linear ($h(g) = \gamma g$), and differencing equation 6 one obtains:

$$\begin{aligned} y_{it+a} - y_{it} &= (a\alpha + 1)(y_{it} - y_{it-a}) + a(X_{it} - X_{it-a})\beta + \\ &\quad a\gamma(g_{it} - g_{it-a}) + ak(g_{it} - g_{it-a})) + ak(g_{it-a} - g_{it-2a})) + a\epsilon_{it} - a\epsilon_{it-a} \end{aligned}$$

or:

$$y_{it+a} - y_{it} = (a\alpha+1)(y_{it} - y_{it-a}) + a(X_{it} - X_{it-a})\beta + a\phi(g_{it} - g_{it-a})) + k(g_{it-a} - g_{it-2a})) + a\epsilon_{it} - a\epsilon_{it-a} \tag{12}$$

where $\phi(x) = k(x) + \gamma x$.

In principle, this equation could be estimated. Using methods similar to those derived in Porter (1996), one could also recover $k(\cdot)$ and γ, but the data requirement would make the exercise senseless in the present context (there are too few countries with three successive measures of inequality).

However, if equation 12 is indeed the correct way to represent the relationship between changes in inequality and growth in the first differenced equation, it suggests that the interpretation of the fixed effects, first difference and GMM estimates of equation 9 could be very misleading. In order to investigate this point without relying on our (potentially biased) estimates of equation 6, we estimate a modified version of equation 11, which does not restrict the coefficient of the difference $g_{it} - g_{it-a}$ to be linear. In other words, we estimate the relationship

$$y_{it+a} - y_{it} = (a\alpha + 1)(y_{it} - y_{it-a}) + a(X_{it} - X_{it-a})\beta + a\phi(g_{it} - g_{it-a}) + a\epsilon_{it} - a\epsilon_{it-a}, \tag{13}$$

where $\phi(\cdot)$ is a function which we want to estimate flexibly. Under the hypothesis that the model in equation 9 is the correct model, we should not be able to reject the linearity.

We use kernel regression, and we "partial out" the linear part of the model using the same methodology we used before. The results are presented in figures 6A and 6B. The linearity seems, once again, to be rejected. To further explore this, we used the same specifications as in Section 3. To test whether the non-linearity is statistically significant, we present in panel C of Table 6 the F test for the joint significance of the non-linear terms in the partially linear model (columns (1) and (2)). Linearity is rejected in both cases, at 10% and 3% levels of confidence, respectively. Panel D presents the results of estimating a quadratic specification for $\phi(\cdot)$. Finally, we estimate a piece-wise linear specification for $\phi(\cdot)$. The coefficients of decreases and increases in inequality are positive and negative respectively. The positive coefficient in the decreasing range is significant. The negative coefficient in the increasing range is smaller in absolute value and insignificant.

To ensure that the non-linearity of the relationship between inequality and growth is not driven by some misspecification in our estimation of the partially linear model,[23] we then test for linearity *under the assumption that the model in the literature we are critiquing was actually correctly estimated.*

To do so, we estimate the main equation 9 using each method, and we then compute:

$$(y_{it+a} - y_{it})^* = y_{it+a} - y_{it} - (a\hat{\alpha} + 1)(y_{it} - y_{it-a}) - a(X_{it} - X_{it-a})\hat{\beta} \tag{14}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the values of α and β obtained by estimating equation 9 by each method. If the assumptions necessary for the validity of each method are satisfied, α and β will be estimated consistently. Then, according to equation 11, the relationship between $(y_{it+a} - y_{it})^*$ and $g_{it} - g_{it-a}$ should be linear.

The next step is to make sure that the estimates of γ obtained if we regress $(y_{it+a} - y_{it})^*/a$ on the difference $(g_{it} - g_{it-a})$ are similar to those obtained using a fixed-effects type estimator. OLS estimates are presented in panel A of Table 6. They are alternative estimates of γ, consistent if equation 11 is correctly specified and if the innovation in inequality is not correlated with the innovation in the error term. Except for the first difference estimate, they are not identical to the estimate of γ reported in Table 5, since they use different estimation methods. However, they are also positive and significant, and their magnitude is similar to that of the fixed effect and Arellano and Bond estimates.

Second, we test the linearity assumption. We start by allowing the coefficient to vary with the sign of the difference $(g_{it} - g_{it-a})$. The results indicate that there is a sharp non-linearity. As before, we find that both increases and decreases in inequality are associated with lower subsequent growth (panel B). This suggests that the conclusions of Forbes (1998), and Li and Zou (1998) are not warranted: There is no evidence in the data that increases in inequality are good for growth. In fact, the bulk of the evidence goes in the opposite direction.

[23]For example, we did not deal with the inconsistency introduced by the lagged endogenous regressor.

In Figure 7, we present a kernel regression of $1/a*(y_{it+a}-y_{it})^*$ on the difference $(g_{it}-g_{it-a})$ for most specifications reported in Table 5. The shape of the curve is similar across specifications, and similar to what we had found when we estimated the partially linear model. We have experimented with a variety of other specifications which we do not report here. The results are always similar. In panel C of Table 6, we report the F statistic of the significance of the non-linear terms in a quartic regression of $1/a * (y_{it+a} - y_{it})^*$ on $(g_{it} - g_{it-a})$. Here also, the data clearly rejects linearity in almost all specifications.

4.3 Consequences for Estimated Coefficients

4.3.1 Random Effects and Fixed Effects

The results suggest that equation 9 is misspecified. Non-linear terms in past changes are omitted in the regression. Since current levels and past changes are correlated, this introduces a bias in the coefficient of inequality when equation 9 is estimated using random effects.

However, the misspecification is accentuated when the equation is estimated in first differences or using fixed effects. The fixed effect estimation imposes a linear structure on the relationship between the deviation of income from its mean across periods and the deviation of gini coefficients from its mean across periods. Since the relationship between growth and inequality is not monotonic in the first difference, it is also not monotonic when period averages are taken out. The fixed-effect estimator is therefore a weighted average of negative and positive coefficients, which can be positive if the weight given to positive coefficients is bigger. The histogram of changes in the gini coefficient, in Figure 8, combined with Figure 6, shows why the first difference coefficient shown in panel B is positive. The majority of the data points are in the region where changes are positively correlated with growth.

4.3.2 Estimation Using the Arellano and Bond Technique

The Arellano and Bond estimator uses lagged levels of inequality to instrument for changes in inequality with lags. Ignoring longer lags, the reduced form equation implicitly estimated when using the Arellano and Bond technique has the form:

$$(y_{it+a} - y_{it})/a = \lambda y_{it-a} + X_{it-a}\kappa + \delta g_{it-a} + \nu_i + \xi_{it} \tag{15}$$

This reduced form is very similar to the equation we had estimated in Section 3. The only difference is that income levels and the control variables are lagged one period. In columns (3) and (6) of Table 3, we present the coefficient of g_{it-a} in this specification. As before, we find a negative, but insignificant, coefficient.

The Arellano and Bond GMM estimator in effect takes the ratio of the negative reduced form coefficient and the negative coefficient from estimating the effect of the level of inequality on changes in inequality. This naturally leads to the positive coefficient in the "structural" equation. For example, dividing -0.033 (column (1), Table 3) by -0.087 (column (7), Table 3) leads to 0.38, close to the Arellano and Bond coefficient of 0.58 reported in the corresponding column in Table 5. Therefore, the seemingly dramatic difference in results obtained when we use the Arellano and Bond method are in fact a different interpretation of the same reduced form evidence presented in this paper or in , e.g. Barro (1999).

This interpretation of the reduced form is clearly misleading, because equation 11 is mis-specified. The omission of higher order terms in the difference $g_{it} - g_{it-a}$ causes a failure of the exclusion restriction. There is a positive relationship between g_{it-a} and $(g_{it} - g_{it-a})^2$ and, in turn, a negative relationship between $(y_{it+a} - y_{it})^*/a$ and $(g_{it} - g_{it-a})^2$, which is assumed away.

5 Conclusion

The main goal of this paper is to describe the cross-country evidence on inequality and growth. We find that there are a number of stable relationships in the data but they do not fit very well with the linear models that have been used to interpret the data. While there are several possible ways to interpret the data, the more fruitful approaches are the ones which take the inherent non-linearities present in the data seriously.

Drawing policy conclusions from our evidence is, however, not easy, as we share with the rest of the literature the problem that the exogeneity of inequality (or changes in inequality) cannot be taken as given. At best our evidence may be seen as a warning against institutional frameworks and political traditions (such as populism) that generate large swings in the income distribution.

On the more fundamental question of whether inequality is bad for growth, our data has little to say. It is clear that the most compelling evidence on this point has to come from micro data. While some interesting evidence is beginning to trickle in,[24] we are only at the beginning of an enormous enterprise.

References

Acemoglu, Daron, and James Robinson (2000) 'Why did the west extend the franchise? Democracy, inequality and growth in historical perspective.' Forthcoming, Quarterly Journal of Economics

Aghion, Philippe, and Patrick Bolton (1997) 'A trickle-down theory of growth and development with debt overhang.' *Review of Economic Studies* 64(2), 151–72

Alesina, Alberto, and Dani Rodrik (1994) 'Distributive politics and economic growth.' *Quarterly Journal of Economics* 109(2), 465–490

Alesina, Alberto, and Roberto Perotti (1996) 'Income distribution, political instability and investment.' *European Economic Review* 81(5), 1170–1189

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, 277–297

Atkinson, A.B., and A. Brandolini (1999) 'Promise and pitfalls in the use of "secondary" data-sets: Income inequality in oecd countries.' Mimeo, Nuffield College, Oxford

Banerjee, Abhijit, and Andrew Newman (1993) 'Occupational choice and the process of development.' *Journal of Political Economy* 101(2), 274–298

Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray (1998) 'Inequality, control rights and rent seeking: A theoretical and empirical analysis of sugar cooperatives in Maharashtra.' Mimeo

[24] For example, Banerjee, Mookherjee, Munshi and Ray (1998) show, using a panel of data from sugar cooperatives in India, that the most unequal cooperatives (in terms of land ownership among cooperative members) are the least productive, with a difference of more than 50% (measured by capacity, which is a proxy for output) between the most and least egalitarian cooperatives.

Bardhan, Pranab, Samuel Bowles, and Herbert Gintis (2000) 'Wealth inequality, wealth constraints and economic performance.' In *Handbook of Income Distribution,* ed. A.B. Atkinson and F. Bourguignon

Barro, Robert J. (1999) 'Inequality, growth, and investment.' Working Paper 7038, National Bureau of Economic Research, March

Benabou, Roland (1996a) 'Inequality and growth.' In *NBER Macroeconomics Annual 1996,* ed. Ben Bernanke and Julio J. Rotemberg (MIT Press) pp. 11–73

—— (1996b) 'Unequal societies.' New York University C.V. Starr Center Economic Research Report: 9617

—— (1999) 'Unequal socities: Income distribution and the social contract.' Mimeo, forthcoming *American Economic Review*

Benhabib, Jess, and A. Rustichini (1998) 'Social conflict and growth.' *Journal of Economic Growth* 1(1), 143–158

Benhabib, Jess, and Mark M. Spiegel (1998) 'Cross-country growth regressions.' Mimeo

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, 363–389

Deaton, Angus, and Christina Paxson (1998) 'Aging, saving and the welfare of the elderly in developing Asia.' Mimeo, Princeton University

Deininger, Klaus, and Lyn Squire (1996) 'A new data set measuring income inequality.' *World Bank Economic Review* 10, 565–591

Forbes, Kristin J. (1998) 'A reassessment of the relationship between inequality and growth.' Mimeo

Galor, Oded, and Joseph Zeira (1993) 'Income distribution and macroeconomics.' *Review of Macroeconomic Studies* 60, 35–52

Hausman, Jerry A., and Whitney Newey (1995) 'Nonparametric estimation of exact consumers surplus and deadweight loss.' *Econometrica* 63(6), 1445–1476

Krueger, Alan, and Mikael Lindahl (1999) 'Education for growth: Why and for whom?' Mimeo

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

Loury, Glenn C. (1981) 'Intergenerational transfers and the distribution of earnings.' *Econometrica* 49(4), 843–867

Perotti, Roberto (1996) 'Growth, income distribution and democracy.' *Journal of Economic Growth* 1, 149–187

Persson, Torsten, and Guido Tabellini (1991) 'Is inequality harmful for growth? Theory and evidence.' *American Economic Review* 48, 600–621

Piketty, Thomas (1997) 'The dynamics of the wealth distribution and the interest rate with credit rationing.' *The Review of Economic Studies* 64(2), 173–189

Porter, Jack (1996) 'Essays in econometrics.' PhD dissertation, MIT

Robinson, James (1988) 'Root-n consistent semiparametric regressions.' *Econometrica* 56, 931–954

Figure 1-A: Relationship between income growth and lagged gini growth: partially linear model (Perotti variables)

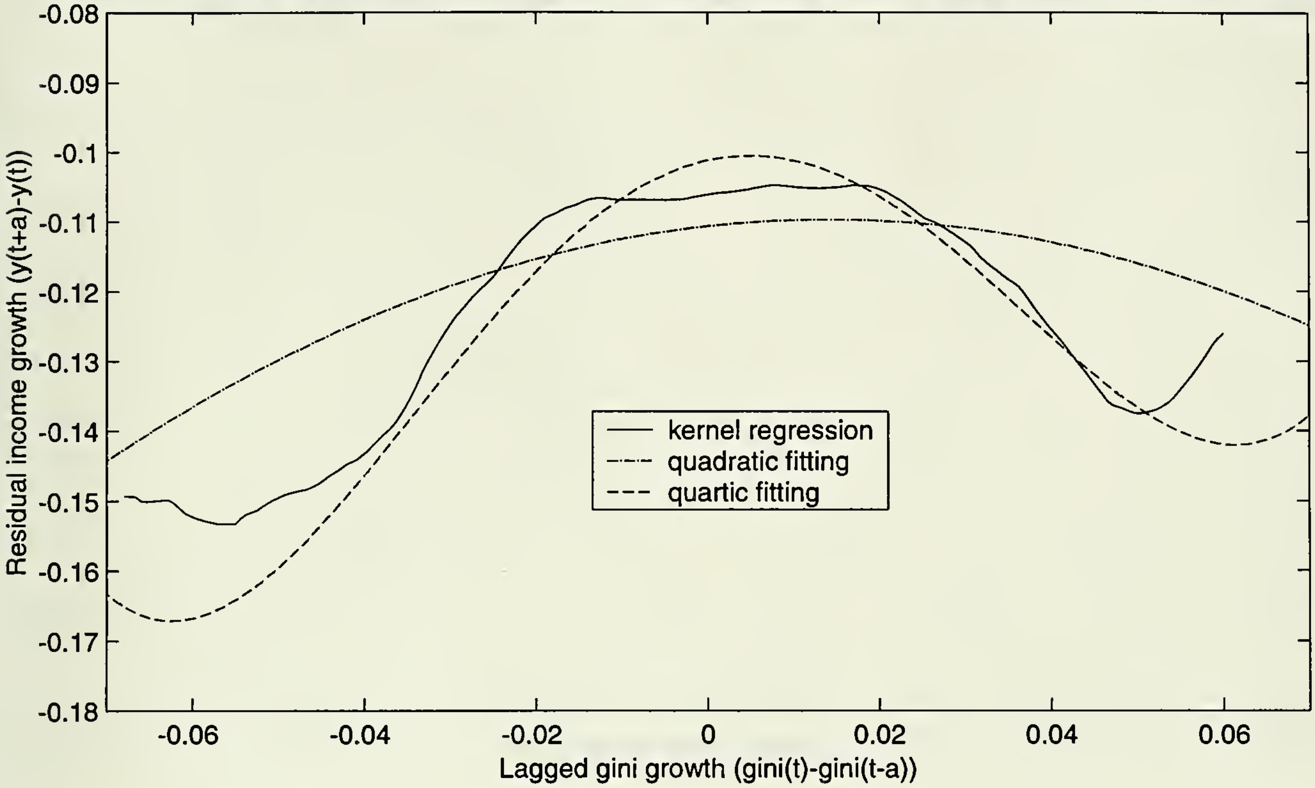

Figure 1-B: Relationship between income growth and lagged gini growth: partially linear model (Barro variables)

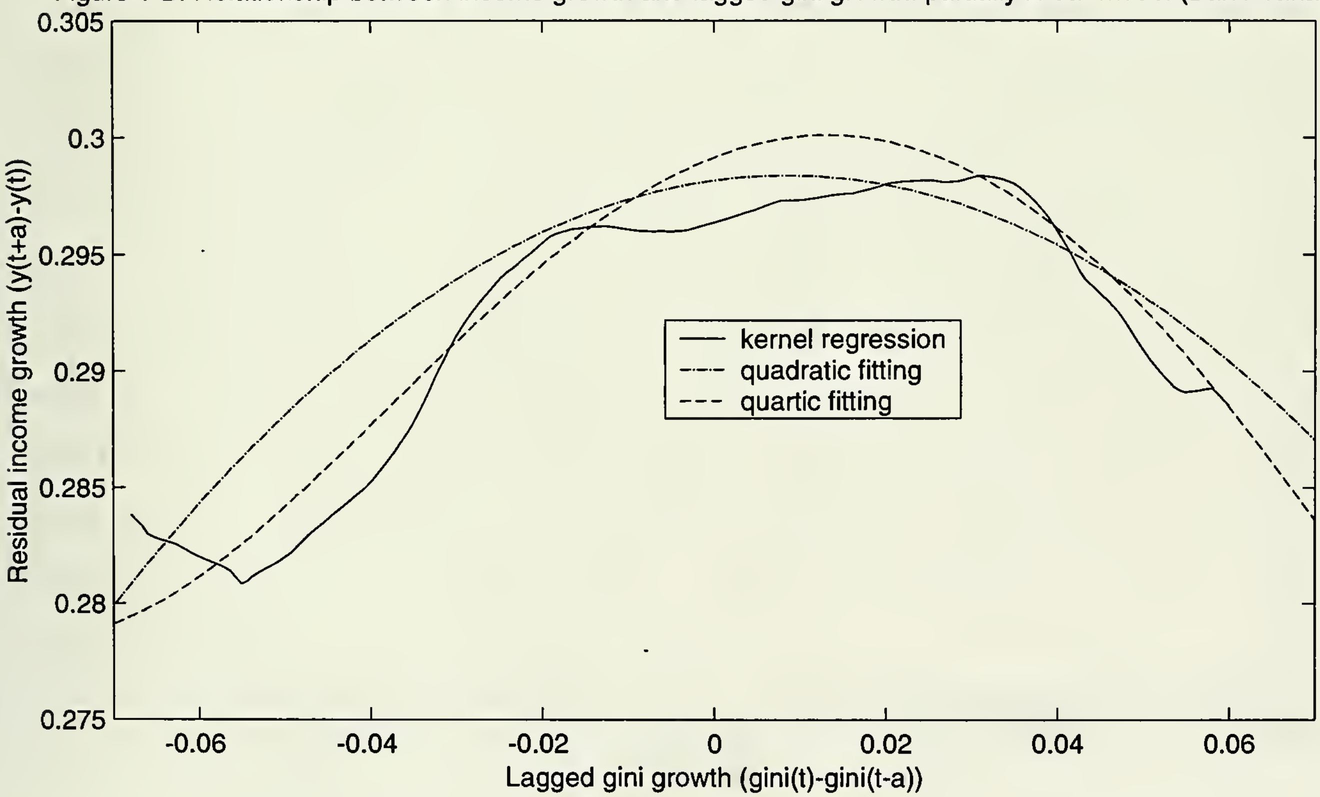

Figure 2: Relationship between income growth and lagged gini: partially linear model (Perotti variables)

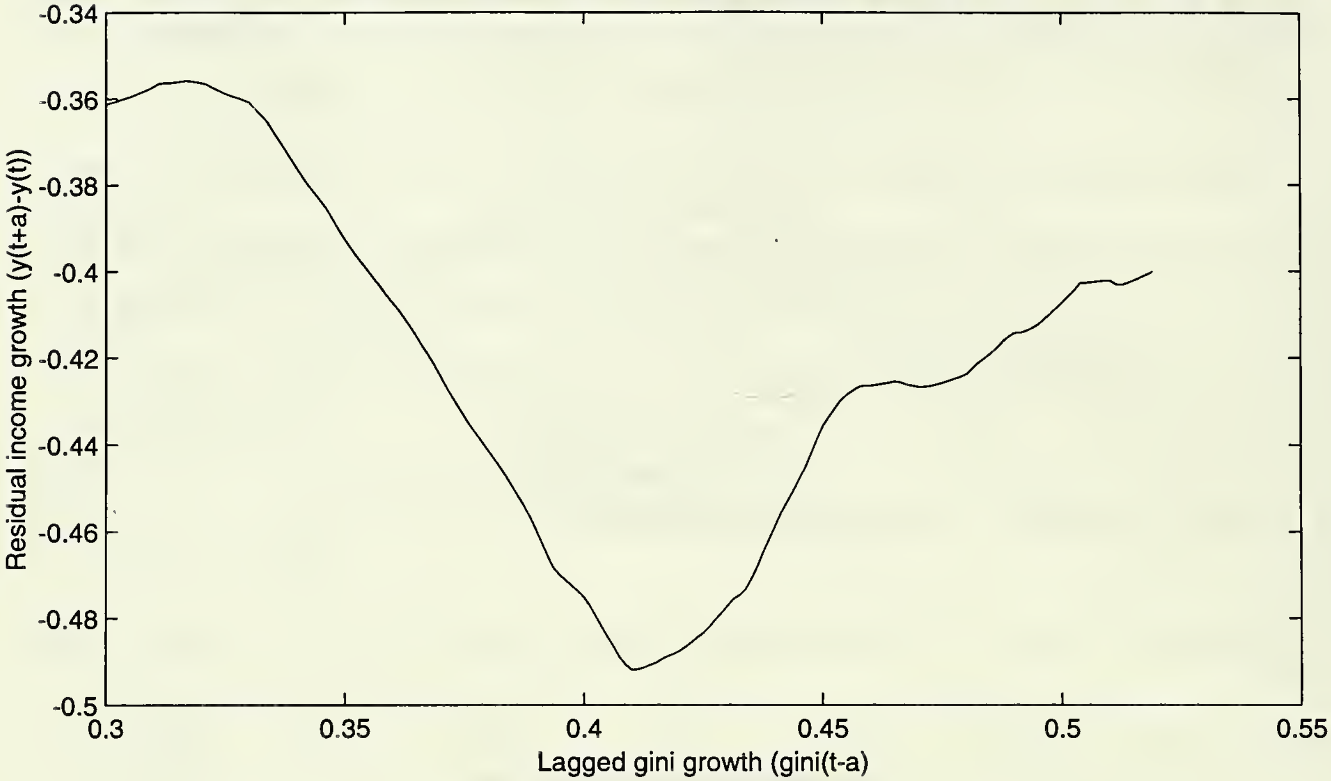

Figure 3: Barro structural form

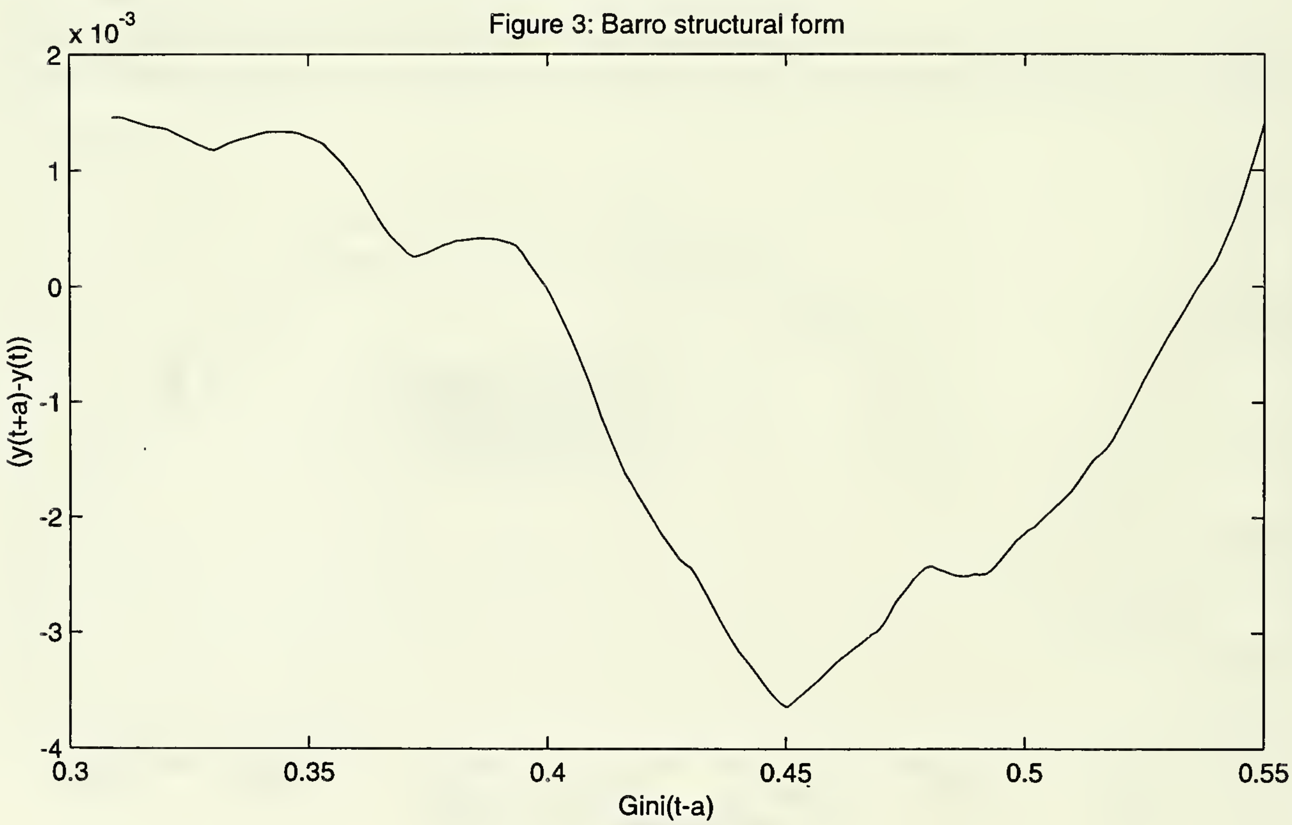

Figure 4A: Barro structural form-Non latin America countries

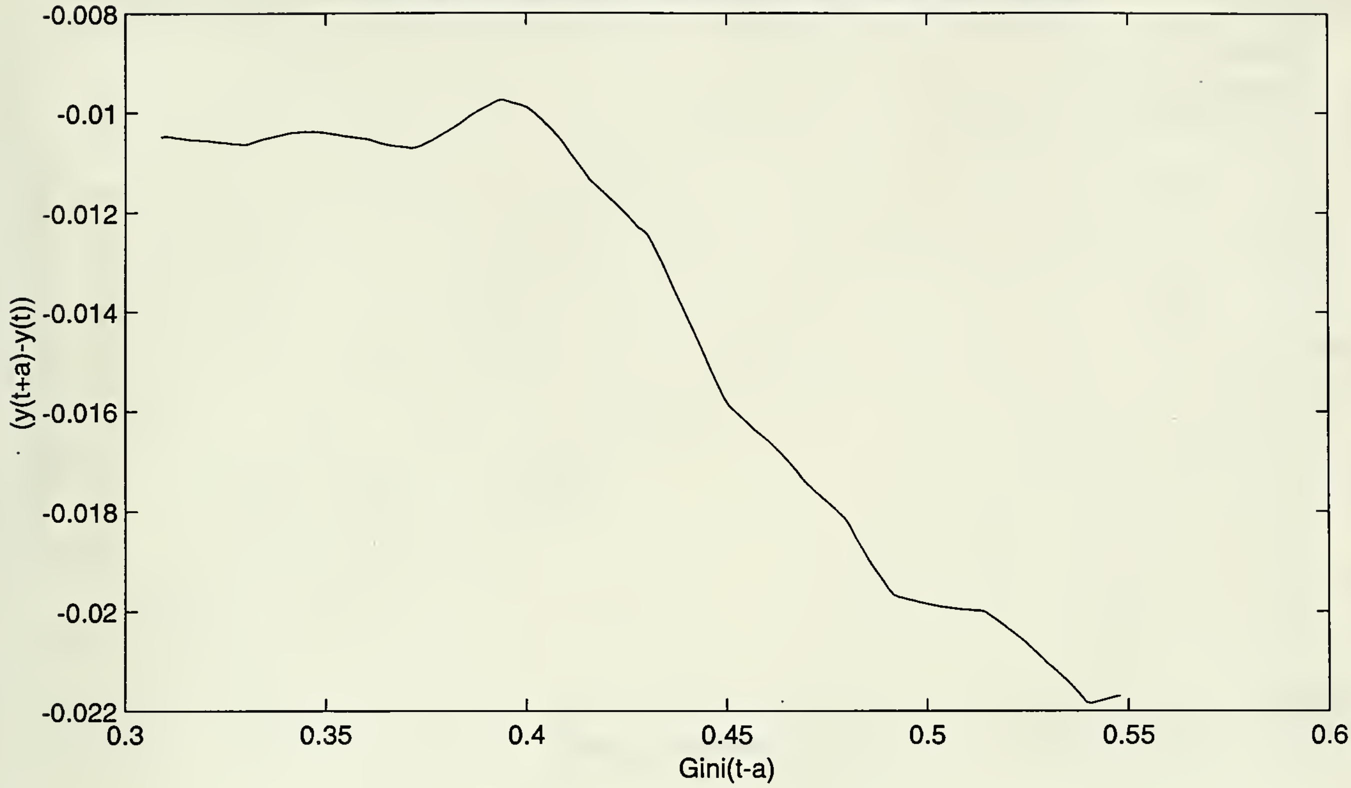

Figure 4B: Barro structural form-Latin American countries

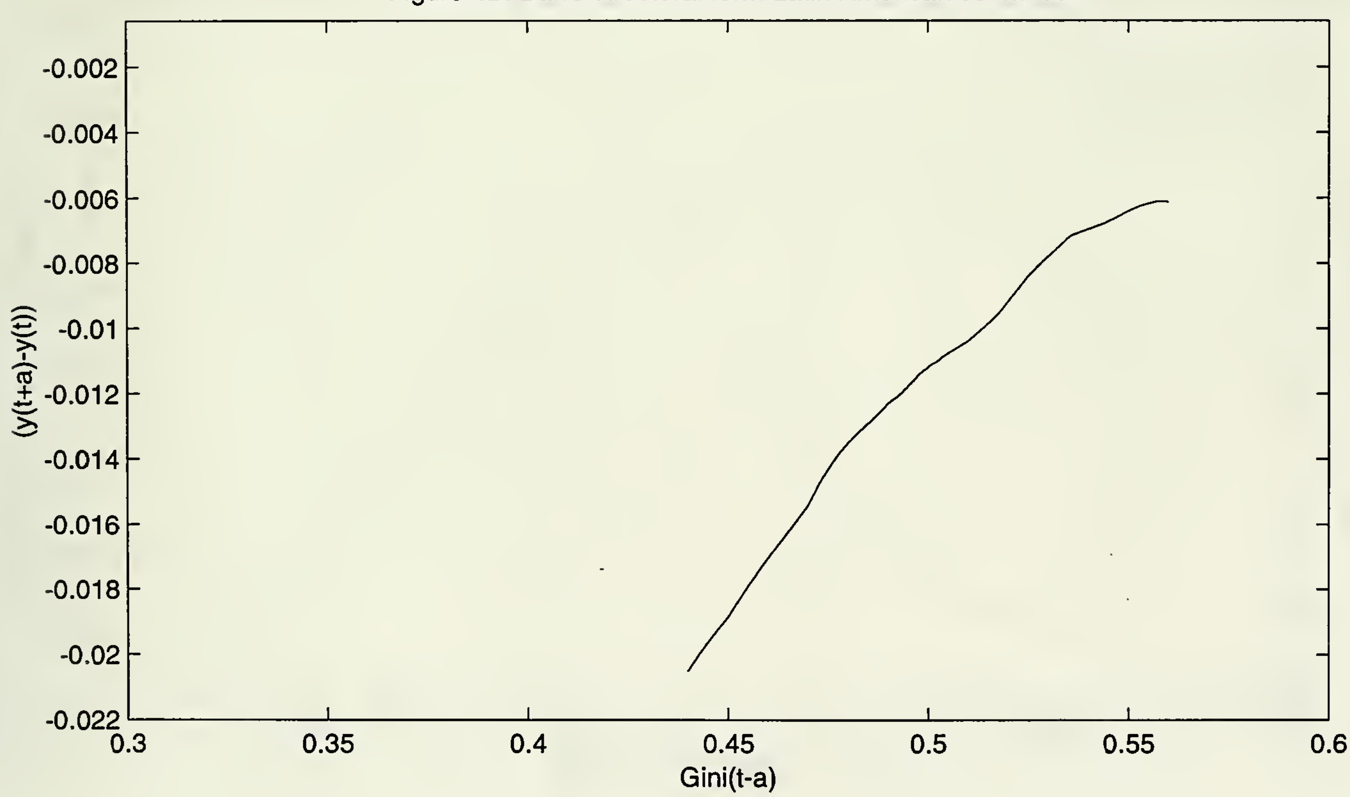

Figure 5A: Relationship between gini growth and lagged gini

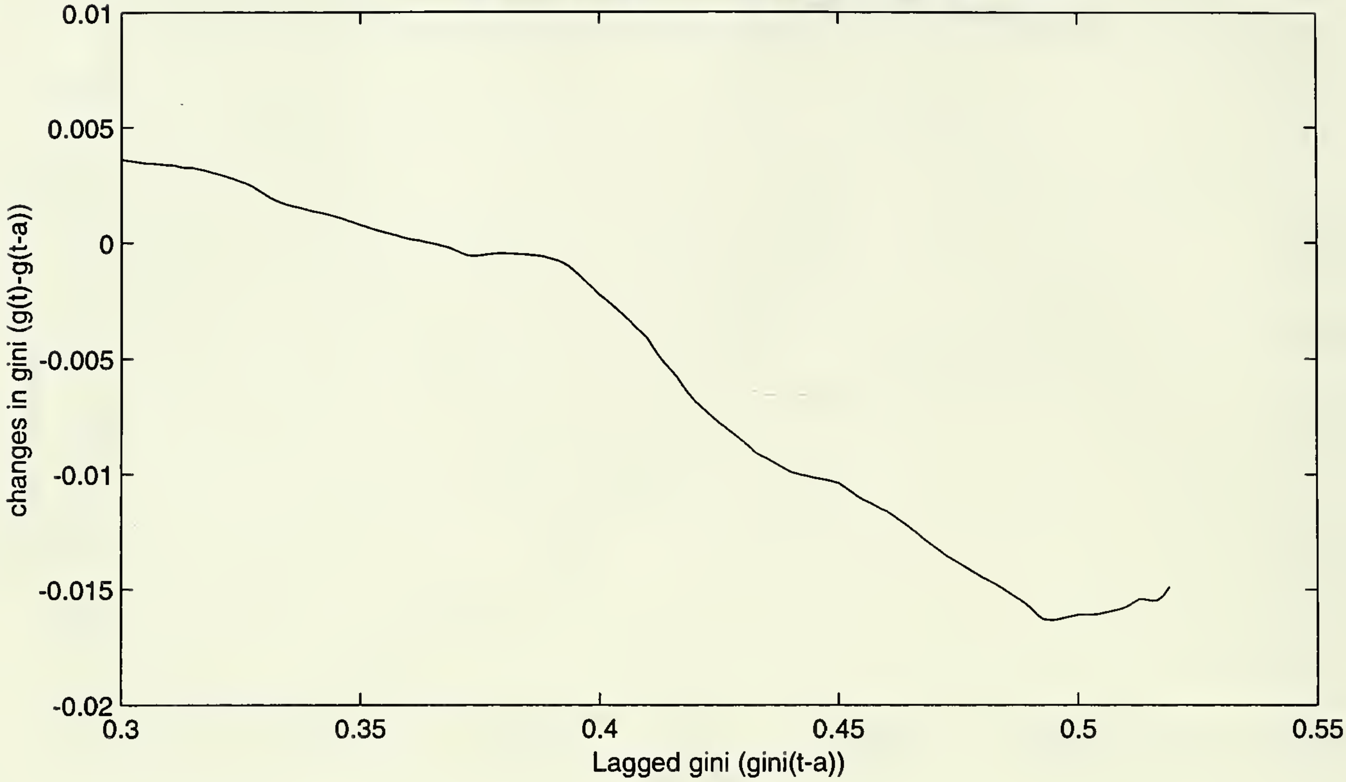

Figure 5B: Relationship between gini growth and lagged gini

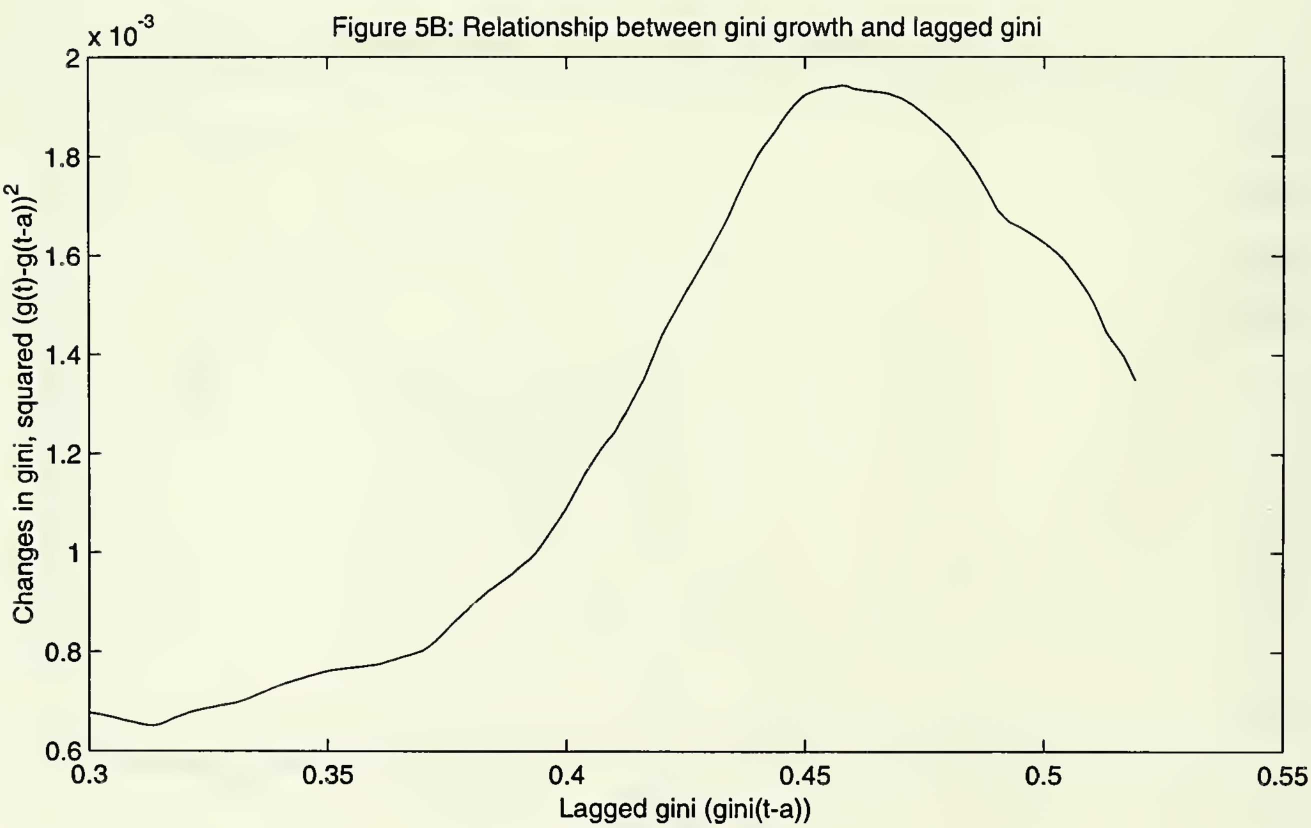

Figure 6-A: Relationship between income growth and lagged gini growth: partially linear model (Perotti variables)

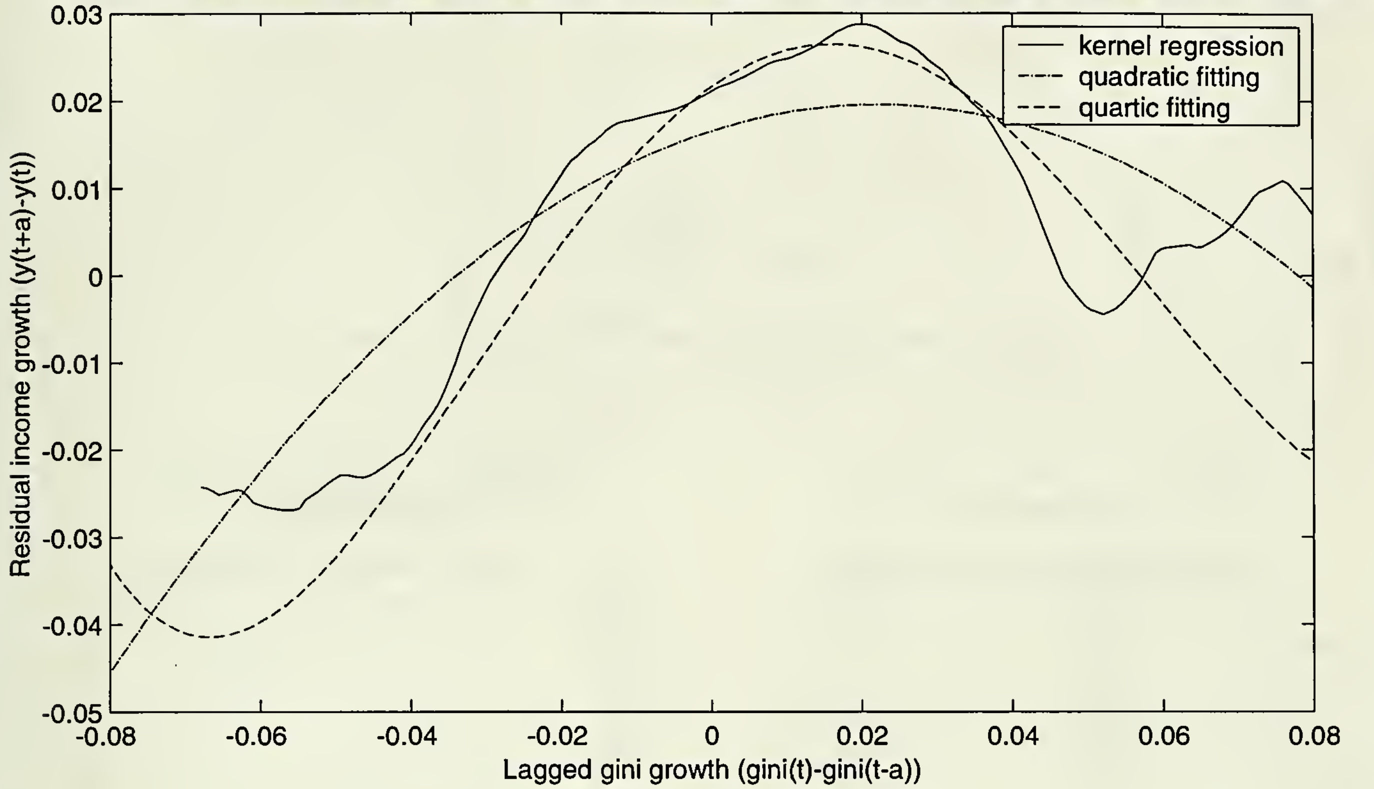

Figure 6-B: Relationship between income growth and lagged gini growth: partially linear model (Barro variables)

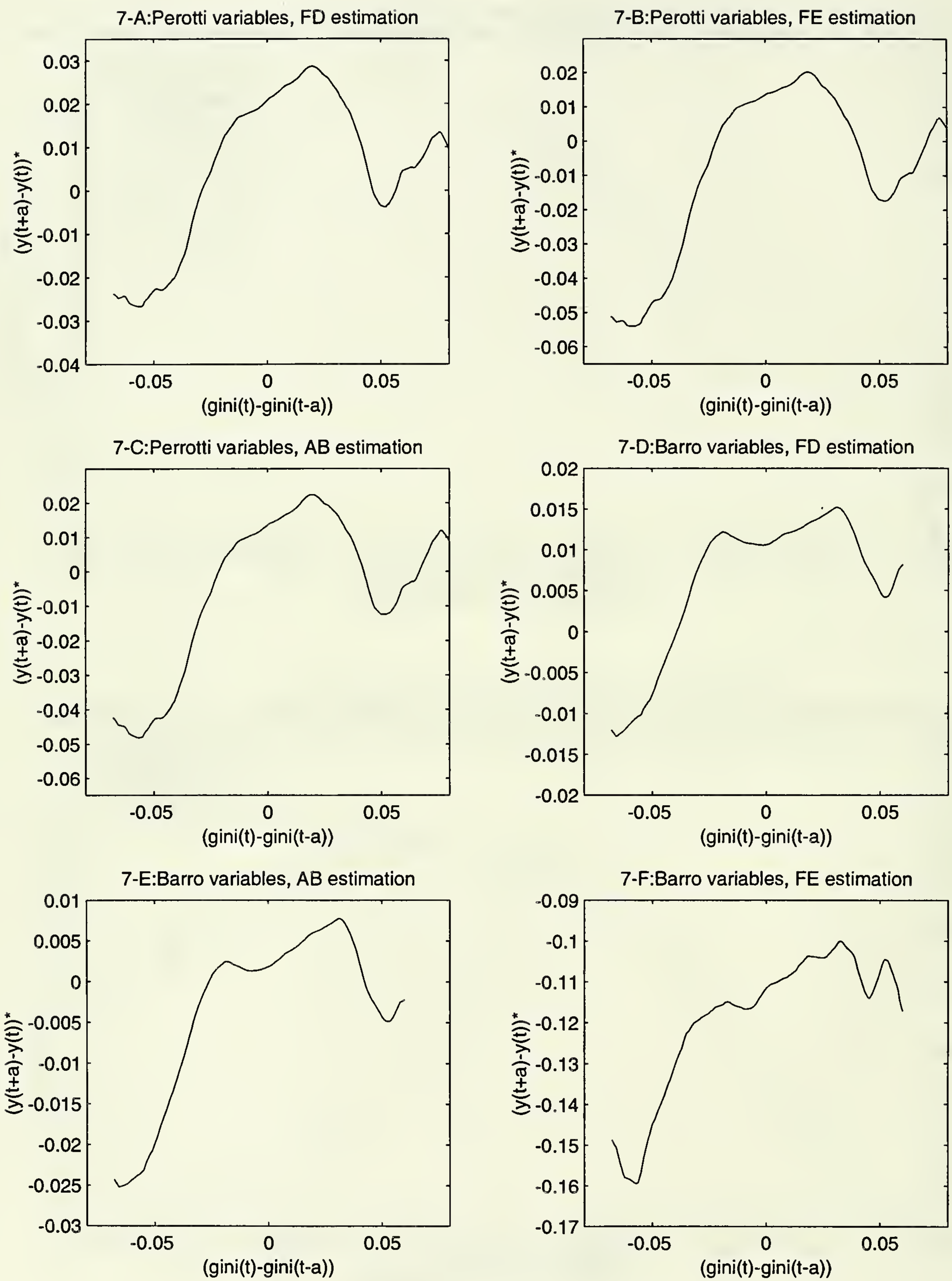
7-A:Perotti variables, FD estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))
7-B:Perotti variables, FE estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))
7-C:Perrotti variables, AB estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))
7-D:Barro variables, FD estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))
7-E:Barro variables, AB estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))
7-F:Barro variables, FE estimation
(y(t+a)-y(t))*
(gini(t)-gini(t-a))

Table 1
Selected descriptive statistic

		D&S, high quality (1)	Barro (2)
A. Mean (standard deviation)			
Log(gdp per capita) in 1980 dollars (Summers and Heston)			
	1965	8.03	7.66
		(0.86)	(0.96)
	1975	8.37	7.95
		(0.85)	(1.01)
	1985	8.58	8.21
		(0.82)	(0.96)
	1995	8.82	8.45
		(0.79)	(1.06)
Gini coefficient (Barro's construction from D&S)			
	1960	0.41	0.43
		(0.077)	(0.06)
	1970	0.39	0.43
		(0.086)	(0.062)
	1980	0.37	0.39
		(0.084)	(0.064)
Gini coefficient (Forbes's construction from D&S)			
	1965		0.38
	1970		0.4
	1975		0.4
	1980		0.38
	1985		0.37
	1990		0.38
B. Number of countries			
		D&S high quality	Barro, 1960
East Asia		7	6
Latin America		9	11
OECD		20	12
Sub-Saharan Africa		0	12
Other		9	9
total		45	50

Table 2

Relationship between inequality and changes in inequality and growth

	Dependent variable: 1/a(y(t)-y(t-a))											
	Perotti specification						Barro specification					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gini (t)	0.021	-0.27		0.05	0.064	0.094	-0.03	-0.29		-0.042		-0.039
	(0.09)	(0.63)		(0.10)	(0.099)	(0.11)	(0.043)	(0.22)		(0.045)		(0.043)
gini (t)2		0.0036						0.33				
		(0.0077)						(0.27)				
gini(t)-gini(t-a)			0.065	0.36					0.053	0.073		
			(0.16)	(0.17)					(0.063)	(0.066)		
(gini(t)-gini(t-a))2			-5.09	-5.37					-2.47	-2.33		
			(2.95)	(3.06)					(1.16)	(1.17)		
gini(t)-gini(t-a)* 1(gini(t)-gini(t-a))<0					0.63						0.27	
					(0.30)						(0.10)	
gini(t)-gini(t-a)* 1(gini(t)-gini(t-a>)=0					-0.59						-0.11	
					(0.33)						(0.13)	
F test for (gini(t)-gini(t-a))2, (gini(t)-gini(t-a))3, (gini(t)-gini(t-a))4						9.02						5.72
(p value in parentheses)						(0.029)						(0.12)
Number of observations	128	128	128	128	128	128	98	98	98	98	98	98

Note: Coefficient obtained using random effect specifications.

Standard errors in parentheses; a is equal to 5 (Five-year periods)

Control variables:

In Perotti specification : Log(GDP(t),PPP I (t), male education (t), female education (t)

In Barro's specification Log(GDP(t-1)), log(GDP(t-1)) squared, government consumption(t-1), secondary education(t), higher education(t), fertility(t), (term of trade(t+1)-terms of trade(t)), rule of law, democ(t), democt(t) squared, spanish or portuguese colony,other colony, investisment share (t-1)

Table 3

Estimation of the reduced form model

	Dependent variable: (y(t+a)-y(t))/a						Dependent variable: change in gini coefficient			
	Perotti			Barro			Perotti		Barro	
							g(t)-g(t-a)	$(g(t)-g(t-a))^2$	g(t)-g(t-a)	$(g(t)-g(t-a))^2$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
g(t-a)	-0.047	0.77	-0.033	-0.043	-0.21	-0.10	-0.087	0.0067	-0.25	0.0076
	(0.076)	(0.66)	(0.082)	(0.039)	(0.21)	(0.043)	(0.038)	(0.0025)	(0.066)	(0.0038)
g(t-a)^2		-0.94			0.26					
		(0.81)			(0.27)					
Control variables	X(t)	X(t)	X(t-a)	X(t)	X(t)	X(t-a)	X(t-a)	X(t-a)	X(t-a)	X(t-a)

Note: Coefficient obtained using random effect specifications.

Standard errors in parentheses; a is equal to 5 (Five-year periods)

Control variables: X(t) stands for control variable not lagged.

X(t-a) stands for control variables lagged one period (5 years).

For a list of control variables see note to table 2.

Table 4

Estimation in Barro's sample

	Dependent variable: 1/2a*[ln(gdp(t+2a))-ln(gdp(t)]							
	Barro reduced form		Structural model		Structural model		Reduced form	
	Perotti	Barro	Barro		Barro,excluding fertility		Barro	Perotti
	SUR	SUR	3SLS	3SLS	3SLS	3SLS	SUR	SUR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini(t-a)	-0.026	-0.023	-0.00012		-0.037			
	(0.021)	(0.022)	(0.018)		(0.017)			
Gini(t-a)				-0.034		-0.035	-0.057	-0.055
*(1-Latin America Dummy)				(0.020)		(0.025)	(0.023)	-0.026
Gini(t-a)				0.099		0.083	0.11	0.13
*Latin America Dummy				(0.034)		(0.044)	(0.041)	-0.5
Observations	155	146	146	146	146	146	146	158

Notes:

Standard errors in parentheses; a is equal to 5 (Growth is averaged over 10 year periods)

A list of countries in the sample is given in table A1

A list of the control variables is given in the footnote to table 2.

Table 5
Relationship between growth and changes in Gini estimated by differencing methods

	Dependent variable: (y(t+a)-y(t))/a					
	Perotti Specification			Barro Specification		
	First Difference	Fixed effect	Arellano & Bond	First Difference	Fixed effect	Arellano & Bond
	(1)	(2)	(3)	(4)	(5)	(6)
Gini(t)	0.298	0.297	0.56	0.158	0.155	0.27
	(0.18)	(0.16)	(0.039)	(0.068)	(0.063)	(0.016)

Note:
Standard errors in parentheses; a is equal to 5 (Five-year periods)
For a list of control variables see note to table 2.

Table 6

Non-linearity of the relationship between change in gini and growth in models based on first differences

	Dependent variable (y(t+a)-y(t))/a		Dependent variable: residual growth (1/a*[y(t+a)-y(gdp(t)]*)							
Control variables	Perotti	Barro	Perotti				Barro			
Speficication	First difference	First difference	Random effect	First difference	Fixed effect	Arellano & Bond	Random effect	First difference	Fixed effect	Arellano & Bond
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)	(10)	(11)
A. Linear assumption: OLS Coefficient of (Gini(t)-Gini(t-a))										
Gini(t)-Gini(t-a)	0.298	0.158	0.43	0.298	0.37	0.36	0.11	0.158	0.12	0.17
	(0.18)	(0.068)	(0.19)	(0.18)	(0.19)	(0.18)	(0.34)	(0.068)	(0.35)	(0.07)
B.Piecewise linear assumption:OLS coefficients of (Gini(t)-Gini(t-a))										
if Gini(t)-Gini(t-a) <=0	0.79	0.39	0.94	0.61	0.84	0.6985	0.38	0.42	0.42	0.4
	(0.30)	(0.13)	(0.48)	(0.38)	(0.40)	(0.3848)	(0.61)	(0.13)	(0.48)	(0.13)
if (Gini(t)-Gini(t-a))>=0	-0.3	-0.13	-0.57	-0.48	-0.58	-0.49	-1.18	-0.64	-0.44	-0.11
	(0.35)	(0.11)	(0.54)	(0.35)	(0.40)	(0.38)	(0.88)	(0.14)	(0.55)	(0.14)
C.quartic specification										
F-test for non linear terms joinly significant	2.21	3.37	2.41	1.84	3.04	2.55	1.98	2.61	1.55	3.3
	(0.09)	(0.02)	(0.07)	(0.14)	(0.031)	(0.059)	(0.10)	(0.055)	(0.20)	(0.02)
D.quadratic specification										
Gini(t)-Gini(t-a)	0.23	0.13	0.366	0.24	0.298	0.311	0.16	0.13	0.2	0.15
	(0.18)	(0.067)	(0.24)	(0.17)	(0.19)	(0.19)	(0.35)	(0.063)	(0.24)	(0.66)
$(Gini(t)-Gini(t-a))^2$	5.88	-3.24	-6.94	-5.59	-7.08	-5.94	-8.34	-2.94	-4.49	-3.28
	(3.39)	(1.26)	(4.52)	(0.17)	(3.58)	(3.43)	(6.44)	(1.15)	(4.47)	(1.23)
Numer of observations	125	98	125	125	125	125	98	98	98	98

Note:

Standard errors in parentheses; a is equal to 5 (Five-year periods)

For a list of control variables see note to table 2.

For a definition of residual growth, see the text

Table A1

List of countries in the sample

Deininiger and Squire, high quality	Barro 1960	1970	1980
Australia	Denmark	Australia	Australia
Bangladesh	Finland	Fiji	Belgium
Belgium	France	Finland	Denmark
Brazil	Greece	France	Fiji
Bulgaria	Hungary	Greece	Finland
Canada	Netherlands	Hungary	France
Chile	Norway	Ireland	Greece
China	Sri Lanka (Ceylon)	Italy	Hungary
Colombia	Sweden	New Zealand	Ireland
Costa Rica	Taiwan	Norway	Italy
Denmark	Thailand	Portugal	Luxembourg
Dominican Republic	U.K.	Singapore	Netherlands
Finland	W. Germany	Spain	New Zealand
France	Yugoslavia	Sri Lanka (Ceylon)	Norway
Germany	Argentina	Sweden	Poland
Greece	Bangladesh (E. Pakistan)	Taiwan	Portugal
Hong Kong	Benin	Thailand	Singapore
Hungary	Brazil	Turkey	Spain
India	Burma (Myanmar)	U.K.	Sri Lanka (Ceylon)
Indonesia	Canada	W. Germany	Sweden
Ireland	Chad	Yugoslavia	Switzerland
Italy	Colombia	Argentina	Taiwan
Japan	Costa Rica	Bahamas	Thailand
Korea, Republic of	Egypt	Bangladesh (E. Pakistan)	U.K.
Malaysia	Gabon	Barbados	W. Germany
Mexico	Guyana	Bolivia	Yugoslavia
Netherlands	India	Brazil	Bahamas
New Zealand	Indonesia	Canada	Bangladesh (E. Pakistan)
Norway	Iraq	Chile	Barbados
Pakistan	Israel	Colombia	Brazil
Peru	Ivory Coast (Cote)	Costa Rica	Cameroon
Philippines	Jamaica	Ecuador	Canada
Poland	Japan	Egypt	Chile
Portugal	Kenya	El Salvador	China
Singapore	Madagascar	Gabon	Colombia
Spain	Mexico	Honduras	Costa Rica
Sri Lanka	Morocco	Hong Kong	Dominican Rep.
Sweden	Niger	India	El Salvador
Thailand	Nigeria	Indonesia	Gabon
Trinidad and Tobago	Peru	Iran	Guatemala
Tunisia	Philippines	Jamaica	Hong Kong
Turkey	S. Korea	Japan	India
United Kingdom	Senegal	Liberia	Indonesia
United States	Suriname	Malawi	Iran
Venezuela	Tanzania	Malaysia	Ivory Coast
	Togo	Mexico	Japan
	Trinidad &Tobago	Pakistan	Jordan
	U.S.A.	Panama	Kenya
	Venezuela	Peru	Malawi
	Zambia (N. Rhodesia)	Philippines	Malaysia
		S. Korea	Mauritius
		Senegal	Mexico
		Sierra Leone	Morocco
		Sudan	Nepal
		Tanzania	Nigeria
		Trinidad & Tobago	Pakistan
		Tunisia	Panama
		Uganda	Peru
		Uruguay	Rwanda
		U.S.A.	S. Korea
		Venezuela	Seychelles
		Zambia (N. Rhodesia)	Sierra Leone
			Tanzania
			Trinidad & Tobago
			Tunisia
			U.S.A.
			Venezuela
			Zambia (N. Rhodesia)

Table A2

Countries with large changes in gini coefficients

Decrease in gini coefficient larger than 3 percentage points				Increase in gini coefficient larger than 3 percentage points			
Country	Period	Begining of period gini (in %)	Change in gini (percentage points)	Country	Period	Begining of period gini (in %)	Change in gini (percentage points)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bangladesh	65-70	37.3	-3.1	Australia	85-90	37.6	4.1
Bulgaria	70-75	21.5	-3.7	Bulgaria	75-80	17.8	7.2
Brazil	75-80	61.9	-4.2	Brazil	80-85	57.8	4.0
Canada	85-90	32.8	-5.3	Brazil	70-75	57.6	4.3
Colombia	70-75	52.0	-6.0	Chile	75-80	46.0	7.2
Spain	75-80	37.1	-3.7	China	85-90	31.4	3.2
Finland	70-75	31.8	-4.8	Colombia	75-80	46.0	8.5
Finland	85-90	30.8	-4.7	Germany	65-70	28.1	5.4
France	75-80	43.0	-8.1	Dominican Republic	85-90	43.3	7.2
Hong Kong	85-90	45.2	-3.2	Finland	75-80	27.0	3.9
Hungary	65-70	25.9	-3.0	United Kingdom	85-90	27.1	5.2
Indonesia	80-85	42.2	-3.2	Hong Kong	80-85	37.3	7.9
Ireland	75-80	38.7	-3.0	Sri Lanka	75-80	35.3	6.7
Italy	75-80	39.0	-4.7	Sri Lanka	80-85	42.0	3.3
Korea, Republic	80-85	38.6	-4.1	Mexico	85-90	50.6	4.4
Sri Lanka	85-90	45.3	-8.6	New Zealand	85-90	35.8	4.4
Sri Lanka	65-70	47.0	-9.3	New Zealand	75-80	30.0	4.8
Mexico	75-80	57.9	-7.9	Sweden	75-80	27.3	5.1
Norway	75-80	37.5	-6.3	Thailand	85-90	43.1	5.7
Portugal	75-80	40.6	-3.8	Venezuela	80-85	39.4	3.4
Sweden	70-75	0.4	-6.1	Venezuela	85-90	42.8	11.0
Trinidad and To	75-80	51.0	-4.9				
Trinidad and To	80-85	46.1	-4.4				
Turkey	70-75	56.0	-5.0				
Venezuela	75-80	47.7	-8.2				

Deininger and Squire high quality sample