cef.up
Centro de Economia e Finanças da UPorto

CEF.UP Working Paper 2012-04

A META-ANALYTIC ASSESSMENT OF THE EFFECTS OF INEQUALITY ON GROWTH

Pedro Neves,
Sandra Silva,
and
Óscar Afonso

Centro de Economia e Finanças da UP
www.fep.up.pt

A meta-analytic assessment of the effects of inequality on growth

Pedro Neves,* Sandra Silva,† Óscar Afonso‡

December 12, 2012

Abstract

Over the last two decades there has been a growing interest in determining the impact of inequality on growth. The empirical literature has, however, produced controversial results regarding both the signal and the magnitude of such impact. This paper develops a meta-analysis on this literature on an attempt to systematize and explain the diversity in studies' results. We find that most of the heterogeneity is due to differences in studies' methodological characteristics, such as the structure of the data, the sample coverage, the type of distribution, the definition of income, and the estimation technique. These results suggest that there is not one but several underlying effects of inequality on growth, which are likely to differ in their nature and operate in opposing directions. We also find traces of publication bias, as, on the one hand, authors and journals are more willing to report and publish statistically significant results, and, on the other hand, studies' results tend to follow a predictable cycle of fashion and novelty over time.

Keywords: meta-analysis, inequality, economic growth, publication bias.

JEL codes: O4, D3, H2, C21

*CEF.UP and Faculdade de Economia, Universidade do Porto

†CEF.UP and Faculdade de Economia, Universidade do Porto. Corresponding author. Address: R. Roberto Frias, 4200-464 Porto, Portugal; +351225571254 (phone); +351225505050 (fax). e-mail address: sandras@fep.up.pt.

‡CEF.UP, OBEGEF, NIFIP and Faculdade de Economia, Universidade do Porto

1 Introduction

In Neves and Silva (2010), we provided a comprehensive and critical review of the literature on the effects of inequality on growth. After briefly surveying the main theoretical approaches, we conducted an exhaustive study of the empirical works on this research field and made a systematization of their results. Apparently, there exists a lack of consensus in the assessment of the inequality-growth relationship, as the empirical studies present very different results regarding the signal and the magnitude of the relationship and the validity of some of its transmission channels. We then made a summarization of some of the methodological issues at stake and, based on these, derived heuristically some possible explanations for the variety in the results.

Although useful in systematizing the results of this vast literature, finding possible causes for their diversity, and offering guidelines for a better comprehension of the nature of the inequality-growth relationship, such an approach may be subject to criticism for its lack of a quantitative fundamentation. In fact, as stated by Stanley and Jarrell (1989), traditional literature reviews are vulnerable to a great deal of subjectivity, as reviewers may choose which studies to include in the review, what weights to attach to each study, how to interpret the results, and which factors are responsible for the differences among them. Moreover, in a research field marked by a great diversity of findings and methodologies, intelligent summary may be very difficult, hence wrong interpretations and misleading review conclusions may occur frequently. In order to rule out such criticisms and potential problems, we develop a meta-analysis of the empirical literature on the effects of inequality on growth.

Meta-analysis is a quantitative literature review method, in which statistical procedures are used to contrast and combine results from different studies investigating the same research question, with the aim of identifying patterns among them, sources of disagreement, or other interesting relationships that may come to light in the context of multiple studies (Greenland and O'Rourke, 2008). Used initially in medical and psychological research (Rosenthal, 1984; Hunt, 1997; Hedges and Olkin, 1985), meta-analysis has spread to other research fields, and today is used in several social sciences, including economics.

The aim of this paper is then to employ meta-analytic techniques to further investigate the results of the inequality-growth empirical literature, by summarizing and systematizing them in a more objective way. In particular, we address three important questions. First, based on the estimates reported by the empirical studies, we check what the overall effect of inequality on growth is, paying special attention to its direction and magnitude. Second, we evaluate at which extent the results in this field are distorted by publication bias, a problem which has been widely recognized as an important threat to the validity of empirical research. Third, we use econometrics to explain why the results in this literature differ so much, thereby testing the hypotheses advanced in Neves and Silva (2010).

As for the first question, we find that the overall impact of inequality on growth is insignificant, both statistically and economically, which means that on average the relationship between the two variables is weak. In spite of this fact, there seems to be not one but several "true" effects of inequality on growth, which are likely to differ in their nature and operate in opposing directions. This is suggested by the manifest heterogeneity in the reported results. Regarding the second question, the evidence suggests that there are traces of publication bias in this literature, as, on the one hand, authors and journals are more willing to report and publish statistically significant results, and, on the other hand, studies' results tend to follow a predictable cycle of fashion and novelty over time. Concerning the third question, we find that most of the heterogeneity in the reported results is indeed due to the hypotheses advanced in Neves and Silva (2010).

The paper is set out as follows. Section 2 presents briefly the concept of meta-analysis, its objectives and applications. Section 3 provides a description of the studies used in the meta-analysis, and explains in detail the criteria to their inclusion. In Section 4 the meta-analysis itself is conducted. The statistical tools employed to investigate the questions referred in the previous paragraph are presented and the respective results are discussed. Section 5 concludes.

2 Meta-analysis: concept, objectives and applications

Meta-analysis is a quantitative literature review method which has been widely used as an alternative approach to narrative literature review. It is an application of statistical procedures to findings of a set of studies investigating the same question, aimed at attempting to integrate and explain them.

G.V. Glass is usually credited for the development of meta-analysis. In 1976, he succinctly described it as "the analysis of analyses" (Glass, 1976). Ever since, thousands of meta-analyses have been produced in several research areas, especially in psychology and medical sciences, where the experimental setup of the research has made its adoption relatively straightforward. Many prominent examples involve clinical trials of new drugs and medical treatments. In economics, meta-analysis has been increasingly used in the last two decades, particularly in those areas where the empirical literature is not consensual. Estimates of the union wage gap and gender wage discrimination (Jarrell and Stanley, 1990, 2004), evaluations of recreation benefits (Smith and Kaoru, 1990; Rosenberg and Loomis, 2000), the spillover effects of multinational corporations (Gorg and Strobl, 2001), environmental impacts identification (Bergh *et al.*, 1997), tests of the Ricardian hypothesis (Stanley, 1998, 2004, 2005), the relationship between freedom and economic growth (Docouliagos, 2005), estimation of the returns to education (Ashenfelter *et al.*, 1999) are just some of the examples of the economics research fields in which meta-analysis has been performed.

In comparison with traditional qualitative literature reviews, meta-analysis has the advantage of summarizing studies' findings in a systematic way, thus reducing the chances of making wrong interpretations and drawing misleading review conclusions (Shadish, 1982). Indeed, empirical and experimental studies typically come to different findings, making intelligent summary difficult. By combining the results of all studies in one statistical analysis, meta-analysis is better positioned to draw more reliable conclusions. Moreover, quantitative methodology is also more helpful than qualitative reviews in highlighting gaps in the extensive literature (Light and Pillemer, 1984): in particular, meta-analysts can conduct many practical tests, such as identifying moderating variables across a large research literature, detecting the interaction between variables and interpreting trends of studies' findings. Usually, these practical tests provide insight into the topic area under investigation and new directions for conducting relevant researches.

In order to explain the variations in the results of the empirical literature on the effects of inequality on growth, we follow Stanley and Jarrell (1989) who suggest making use of a meta-regression analysis. Meta-regression analysis is a type of meta-analysis especially designed to investigate empirical research in economics (Stanley and Jarrell, 1989), which involves estimating a standard regression model:

$$Y_j = \beta_0 + \sum_{k=1}^{K} \beta_k X_{kj} + e_j, \quad j = 1, 2, ..., N \tag{1}$$

where: Y_j is the reported estimate of the phenomenon of interest in study j in a literature comprised of N studies; X_{kj} are meta-independent variables which measure relevant characteristics of the empirical studies (e.g., sample, model specification and estimation techniques) and explain the variation in Y_j across them; β_k are the meta-regression coefficients reflecting the effect of each characteristic on Y_j; and e_j is the meta-regression disturbance term. Equation (1) can then be estimated to quantify the extent to which some of the elements mentioned in Neves and Silva (2010) influence the inequality-growth relationship.

3 The data

In the first stage of the research we conducted a systematic search of the literature on the impact of inequality on growth via electronic sources. We searched the Economic Literature Index (Econlit)[1] for any reference on "*inequality and growth*" and "*distribution and growth*" in the title or in the abstract of articles published in scientific journals. Our search led to a total of 196 results.[2] Since we are interested in collecting estimates of the effect of inequality on growth in empirical articles that use a sample of several countries, pure theoretical articles, case studies, and studies at the national level were excluded. This left us with 29 papers. In order to guarantee comparability of the population under investigation, we also restricted our sample to studies that use the Gini index as a measure of income inequality. The application of this criterion led to the exclusion of only two of the 29 studies.

We then defined the summary variable to meta-analyze as the partial derivative of the average annual growth rate with respect to the Gini coefficient. This is our central measure of "effect size" (Glass, 1976), as it measures the direction and the magnitude of the correlation between inequality and growth. In fact, most empirical studies estimate a regression of the form:

$$g = \alpha_0 + \sum_{m=1}^{M} \alpha_m Z_m + \delta INEQ + u \qquad (2)$$

where: g is the average annual growth rate (usually measured as a dlog of GDP *per capita*); $INEQ$ is a measure of income inequality (usually the Gini coefficient); Z_m is a set of other variables that influence growth; and u is the usual error term. Parameter δ is thus our effect size and its estimate is collected from each empirical study. It is an appropriate measure of the impact of inequality on growth, as it shows the association between the two variables after controlling for other determinants of economic growth. This definition of the effect size imposes another restriction to the inclusion of studies in our sample, namely the fact that these must estimate a linear model linking inequality and growth. This restriction further excluded two studies.

Applying all the above criteria to the primary list of articles, we were left with a total of 25 studies, which are listed in Table 1. Basic regression information was collected from each of them, namely the estimate of the coefficient on inequality, the associated t-statistics and standard errors, as well as the

[1]The Economic Literature Index is an electronic database service, published by the American Economic Association, that provides bibliographical references to a wide range of the economics literature, including journal articles, books, book reviews, collective volume articles, working papers and dissertations.

[2]These results were generated by a search conducted in October, 2012.

sample size.

A frequent problem in conducting a meta-analysis occurs when more than one estimate of the effect size is given in a study. For example, in one single study, Bleaney and Nishiyama (2004) report 28 estimates of the effect of inequality on growth. Should these results be treated separately as "28 studies" or as "one study"? Since there are only 25 papers in our database, treating Bleaney and Nishiyama's paper as equivalent to 28 separate "studies" would clearly give this paper a disproportionate importance. To avoid giving undo weight to a single study or author, we follow Stanley's (2001) principle of choosing one (or very few) estimate from each study.

That being said, another question arises, namely which criteria should be used to choose the underlying estimate from each study. Stanley (2001) and Stanley and Rose (2005) propose using the average estimate, the median estimate, or the estimate preferred by the author. We opted to choose the estimate preferred by the author for two reasons. First, this estimate is the one that the author believes to be the best and takes as a reference in the paper. Second, the use of the average or the median estimate makes the estimation of a meta-regression virtually impossible, as it implies that each observation in equation (1) may present multiple values for a single explanatory variable, X_k.

The "preferred estimate" in each paper was chosen on the basis of the results highlighted by the authors in the abstract and in the conclusion, or, if these were unclear or absent, on the estimates of the baseline regressions. In some papers, more than one "preferred estimate" was presented. In these cases, in order to avoid subjectivity and minimize potential selection bias, we opted to consider in our dataset various "preferred estimates", up to a maximum of three per article.[3] The application of these criteria led to a total of 45 estimates of the impact of inequality on growth, which constitute our meta-sample.

However, the reduction of potential selection bias and the increase in the sample size resulting from considering more than one estimate per study come at a price, namely the introduction of statistical dependence between observations. Following Hunter and Schmidt (2004), a study can be regarded as statistically independent in this context if it uses the same dataset as a previous study but involves different authors, or if the same authors use differet datasets. Therefore, when two estimates are drawn from the same study, albeit resulting from different modelling or estimation techniques, they are likely to be statistically dependent. Fortunately there are econometric procedures to deal properly with this problem, which will be used and described in Section 4.

[3]We defined a maximum of three estimates per article in order to avoid the above mentioned problem of disproportionate importance among studies.

Columns (1)-(4) of Table 1 list the observations that compose our dataset and the estimates of the effect size reported in the primary studies, as well as the respective estimates of the standard errors and t-statistics.[4]

Table 1- List of the studies included in the meta-sample

(1)	(2)	(3)	(4)	(5)	(6)
Study	$\hat{\delta}$	$\hat{\omega}$	**t**	$\hat{\delta}'$	**t**$'$
Alesina and Rodrik (1994)	-0.0358	0.0198	-1.8100	-0.0008	-0.0386
Alesina and Rodrik (1994)	-0.0500	0.0095	-5.2400	-0.0331	-3.4686
Clarke (1995)	-0.0691	0.0276	-2.5000	-0.0201	-0.7286
Perotti (1996)	-0.0700	0.0246	-2.8400	-0.0263	-1.0686
Perotti (1996)	-0.0300	0.0166	-1.8100	-0.0006	-0.0386
Galor and Zang (1997)	-0.0396	0.0179	-2.2110	-0.0079	-0.4396
Deininger and Squire (1998)	-0.0470	0.0168	-2.8000	-0.0173	-1.0286
Deininger and Squire (1998)	-0.0190	0.0200	-0.9500	0.0164	0.8214
Deininger and Squire (1998)	-0.0340	0.0084	-4.0700	-0.0192	-2.2986
Li and Zou (1998)	0.1490	0.0612	2.4360	0.0407	0.6646
Li and Zou (1998)	0.0310	0.0459	0.6750	-0.0504	-1.0964
Li and Zou (1998)	-0.0890	0.0268	-3.3180	-0.0415	-1.5466
Tanninen (1999)	-0.1220	0.0449	-2.7200	-0.0425	-0.9486
Deininger and Olinto (2000)	-0.0111	0.0029	-3.8276	-0.0060	-2.0562
Deininger and Olinto (2000)	0.0033	0.0023	1.4348	-0.0008	-0.3366
Forbes (2000)	0.1300	0.0600	2.1667	0.0237	0.3953
Sylwester (2000)	-0.0700	0.0300	-2.3333	-0.0169	-0.5619
Barro (2000)	0.0001	0.0180	0.0056	-0.0318	-1.7659
Barro (2000)	0.0540	0.0250	2.1600	0.0097	0.3886
Barro (2000)	-0.0330	0.0210	-1.5714	0.0042	0.2000
Castelló, Doménech (2002)	-0.0170	0.0055	-3.0800	-0.0073	-1.3195
Banerjee and Duflo (2003)	0.1550	0.0630	2.4603	0.0434	0.6889
Banerjee and Duflo (2003)	-0.0300	0.0430	-0.6977	0.0462	1.0737
Banerjee and Duflo (2003)	0.1580	0.0680	2.3235	0.0375	0.5521
De la Croix and Doepke (2003)	-0.0300	0.0100	-3.0000	-0.0123	-1.2286
Bleaney and Nishiyama (2004)	0.0280	0.0185	1.5100	-0.0048	-0.2614
Bleaney and Nishiyama (2004)	0.0230	0.0168	1.3700	-0.0067	-0.4014
Knowles (2005)	-0.0170	0.0082	-2.0800	-0.0025	-0.3086
Knowles (2005)	-0.0130	0.0241	-0.5400	0.0296	1.2314
Knowles (2005)	-0.1350	0.0758	-1.7800	-0.0007	-0.0086
Voitchovsky (2005)	-0.0451	0.0615	-0.7332	0.0639	1.0382

[4] Columns (5) and (6) of Table 1 will be identified in subsection 4.3.2.

Bengoa and Robles (2005)	0.0322	0.0943	0.3419	-0.1347	-1.4295
Sarkar (2007)	-0.0110	0.0055	-2.0000	-0.0013	-0.2286
Castelló (2010)	-0.0480	0.0170	-2.8235	-0.0179	-1.0521
Castelló (2010)	-0.0150	0.0140	-1.0714	0.0098	0.7000
Castelló (2010)	-0.0170	0.0260	-0.6538	0.0291	1.1176
Chambers and Krause (2010)	-0.0079	0.0258	-0.3062	0.0378	1.4652
Khalifa and Hag (2010)	-0.1222	0.0746	-1.6381	0.0099	0.1333
Khalifa and Hag (2010)	0.0656	0.0591	1.1100	-0.0391	-0.6614
David and Hopkins (2011)	-0.0737	0.0240	-3.0708	-0.0312	-1.2994
David and Hopkins (2011)	-0.0285	0.0190	-1.5000	0.0052	0.2714
Woo (2011)	-0.0560	0.0230	-2.4348	-0.0153	-0.6634
Woo (2011)	-0.0330	0.0110	-3.0000	-0.0135	-1.2286
Herzer and Vollmer (2011)	-0.0130	0.0037	-3.5100	-0.0064	-1.7386
Herzer and Vollmer (2011)	-0.0130	0.0019	-6.8800	-0.0097	-5.1086

Legend: $\hat{\delta}$: estimate of the effect size; $\hat{\omega}$: estimate of the standard error of $\hat{\delta}$; t: t-statistic associated to δ; $\hat{\delta}'$: publication bias-corrected estimate of the effect size; t': publication bias-corrected t-statistic associated to δ.

Notes: The effect size indicates the impact of an increase of one percentage point in the Gini coefficient (measured on a 0-100% scale) on the average annual growth rate (measured on percentage). For example, the effect of -0.0358 reported by Alesina and Rodrik (1994) implies that an increase in the Gini coefficient in one percentage points results in an estimated decrease of the average annual growth rate of 0.0358 percentage points.

4 Meta-analysis of the effects of inequality on growth

We now use meta-analytical techniques to further characterize our sample and investigate in more detail the empirical results of the effects of inequality on growth. We start by pooling all the effects reported by the primary studies and computing their weighted averages in order to get a first idea of the combined effect - subsection 4.1. We then test for the presence of heterogeneity in the effect sizes - subsection 4.2. Next, an important issue in empirical research is addressed, namely the fact that the results may be influenced by publication bias. We employ several methods to test and correct for the presence of some forms of bias in the inequality-growth literature - subsection 4.3. Taking into account these results, we then examine if the overall effect of inequality on growth is statistically different from zero, or in other words, if on average inequality really has a statistically significant impact on economic growth - subsection 4.4. Finally, meta-regression techniques are employed in order to explain why the results of this empirical literature differ so much across studies, thereby testing some of the conjectures made in Neves and Silva (2010) - subsection 4.5.

4.1 Fixed and random effects estimates of the effect size

A first natural question is what the combined estimate of all studies is that adequately represents the true underlying effect between inequality and growth; that is, what combined estimate of the effect size do we get if we pool the information of the effect size from all studies? We use two widely employed estimators in meta-analysis: the fixed effects estimator and the random effects estimator.[5] Both of them are weighted averages of the effect size estimates reported by the studies, but they differ in their underlying assumptions.

The fixed effects estimator assumes that there is no heterogeneity among study results and that the different magnitude of the estimates is solely due to sampling variation. Statistically, this is equivalent to the hypothesis that all effect sizes are equal, i.e., $\delta_1 = \delta_2 = \delta_3 = ... = \delta_N = \delta$, where δ_j represents the effect size of the j^{th} observation (in our case, $j = 1, 2, 3, ...45$) and δ is the common "true" effect size. When a series of N studies share a common "true" effect size, δ, it is natural to estimate δ by pooling estimates from each study (Hedges and Olkin, 1985). If the sample sizes of the studies differ, then the estimates from the larger studies will be more precise than the estimates from the smaller studies. In this case, it is reasonable to give more weight to the more precise estimates when pooling. Given that the precision of each estimate of the effect size, $\hat{\delta}_j$, can be measured by the inverse of its variance, $1/\omega_j^2$ (see, e.g., Sutton *et al.*, 2000), the fixed effects estimator of δ can be seen as a weighted average of all $\hat{\delta}_j$, with weights given by $\eta_{j,FE} = 1/\hat{\omega}_j^2$.[6]

The random effects estimator, in turn, assumes heterogeneity among study results. According to this formulation, there is not a single "true" effect across studies; instead, each study has its own "true" effect size, randomly drawn from a larger population with a fixed mean and variance. As a consequence, the observed variability in sample estimates of the effect size has two components: one is the sampling error variation, and the other is random variation of the population effect size. Both sources of variation are assumed to be normally distributed, with mean zero and variances ω_j^2 and τ^2 , respectively. As the fixed effects estimator, the random effects estimator is also an inverse-variance weighted estimator of $\hat{\delta}_j$, although the weights are now equal to $\eta_{j,RE} = 1/\left(\hat{\omega}_j^2 + \hat{\tau}^2\right)$, where $\hat{\omega}_j^2$ represents the estimate of

[5] We note that the meaning of the adjectives "fixed" and "random" in meta-analysis is very different from the usual interpretation for panel data models in standard econometrics, because they refer to assumptions about the underlying population effect size, not differing assumptions about the variation across time and regions in panel studies.

[6] $\hat{\omega}_j^2$ stands for the estimate of the variance of $\hat{\delta}_j$ and is collected from the results reported in each study. See in Column (3) of Table 1 the value of $\hat{\omega}_j$ associated to each observation of our meta-sample.

the within-study variance and $\hat{\tau}^2$ the estimate of the between-study variance (see Hedges and Olkin, 1985, for more details).

We have calculated the fixed and the random effect size estimates of the impact of inequality on growth, by averaging the 45 estimates of δ in our sample, using $\eta_{j,FE}$ and $\eta_{j,RE}$ as weights. The fixed effects estimate, $\hat{\delta}_{FE}$, is negative and equals -0.0112. This means that an increase in the Gini coefficient in one percentage point has an estimated negative impact in the average annual growth rate of 0.0112 percentage points. The random effects estimate, $\hat{\delta}_{RE}$, is also negative, equalling -0.0180. Thus, a preliminary finding of our meta-analysis is that inequality seems to influence growth negatively. However, this inference should be taken with caution, as the results of the fixed effects and the random effects estimators should be carefully interpreted. This is explained mainly by three facts.

First, both estimators assume that all observations are independent. As mentioned above, this assumption may not hold in our meta-analysis, since the observations drawn from the same study are likely to be correlated. In this case, the estimators are not efficient. However, this is not a major problem, because, on the one hand, the estimators remain consistent and, on the other hand, multiple measurements is to be preferred in terms of detecting the "true" effect size (Bijmolt and Pieters, 2001).

Second, even if the negative correlation between inequality and growth is statistically significant, one can legitimately question whether or not it is economically meaningful. In fact, both fixed and the random effect estimates suggest that the correlation is not economically meaningful: an estimated effect of -0.0112 (-0.0180) implies that a substantial increase in the Gini coefficient in 10 percentage points reduces the average annual growth rate in only 0.112 (0.180) percentage points. Such small magnitude carries with little practical significance.

Third, it is possible that the primary studies' estimates are influenced by several forms of publication bias, that may distort them. The issue of publication has been generally recognized as a serious threat to the validity of empirical results in many social and medical sciences, and several statistical method have been developed to deal with this problem. In subsection 4.3 we examine how the estimates that compose our sample are affected by the existence of publication bias and employ some techniques designed to correct for it.

4.2 Testing for heterogeneity of effect sizes

We have just seen that the random effects estimator differs from the fixed effects in that it assumes heterogeneity among effect sizes, that is, the studies do not share a common "true" effect size. We can test this hypothesis by performing the so-called Q-test. The Q-test is formally a test of the null hypothesis of homogeneity $H_0 : \delta_1 = \delta_2 = \delta_3 = ... = \delta_N$ versus the alternative that at least one of the effect sizes δ_j differs from the remainder. This test is based on the statistic (Hedges, 1982):

$$Q = \sum_{j=1}^{N} \frac{\left(\hat{\delta}_j - \hat{\delta}_{FE}\right)^2}{\hat{\omega}_j^2} \tag{3}$$

with all notation as before. If all N studies have the same population effect size (i.e., if H_0 is true), then the test statistic Q has an asymptotic chi-squared distribution with $k-1$ degrees of freedom (Hedges, 1982). Thus, if the obtained value of Q exceeds the upper-tail critical value of the chi-square distribution, the null hypothesis of homogeneity of the underlying population effect sizes is rejected.

In our meta-database, the Q-statistic equals 179.2318, which is larger than 60.4801, the 95% critical value of the chi-square distribution with 44 degrees of freedom. Therefore, the hypothesis of the existence of a single "true" effect size is rejected. This conclusion has three important implications. First, the random effects should be preferred to the fixed effects estimator, as the latter relies on an assumption that does not hold. Second, the result of the Q-test implies that there is excess variation in the reported estimates of the effect size that needs to be explained or somehow accommodated. We do so in subsection 4.5 by means of a meta-regression analysis, in which the characteristics of each study correspond to the moderator variables. Third, such excess variation is not only due to sampling error in the original estimations but also due to the existence of different "true" effects of inequality on growth. The meta-regression analysis will allow us to identify what these different "true" effects are.

4.3 Testing and correcting for the presence of publication bias

Before employing meta-regression analysis to explain heterogeneity in effect sizes, one should investigate if and how the studies' estimates are tainted by the existence of publication bias. This issue has received considerable attention in a number of fields, especially in psychology and medicine (see Berg and Berlin, 1988), and also, more recently, in economics. Generally speaking, publication bias refers to a collective label for a set of distortions in the process of reporting of results (Sutton *et al.*, 2000). These distortions have been generally recognized as an important threat to empirical research, as they seriously prevent obtaining reliable estimates of the phenomenon in analysis.

There are several forms of publication bias. First, due to a number of different reasons, authors and journal editors may be interested in publishing results in a certain direction (either positive or negative). Second, they may submit or publish only (or give preference to) statistically significant findings, that is, they may use statistical significance to screen results, leaving aside non-significant findings or studies. These are the two most common forms of publication bias, which have been widely recognized and properly addressed in the literature (Sutton *et al.*, 2000; Stanley, 2005). Other potential forms of bias include: the prior expectations, beliefs and ideological positions of the authors; the type and the editorial positioning of the journals; the country of authors' affiliate institutions; and the existence of predictable patterns of research across time (Doucouliagos *et al.*, 2005; Stanley *et al.*, 2008). In this section we investigate if and how the results of the studies that compose our sample are affected by these categories of bias, by first testing for their presence, and then employing the appropriate statistical methods to accommodate and correct them.

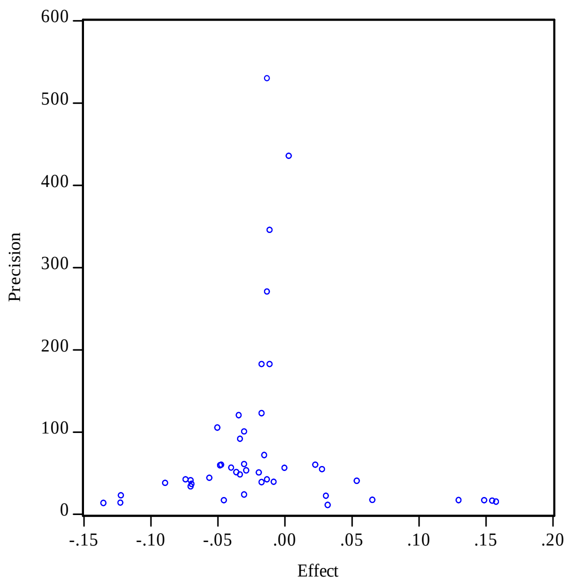
4.3.1 Bias in the direction of the results

One of the most frequent forms of bias in empirical studies is the tendency for authors and journal editors, motivated by, for example, their prior expectations, beliefs or ideological positions, to publish results in a certain direction, either positive or negative. This form of bias is rather pernicious, as it may lead to completely wrong conclusions regarding the nature of the phenomenon in question. For example, Card and Krueger (1995) and Stanley (2005) find that there is a clear bias towards negative findings in the employment effect of minimum wages empirical literature. These findings could suggest that such effect is negative, but after correcting for publication bias the average effect is close to zero. The identification and correction of this type of bias is thus of major relevance, specially when the estimates of the phenomenon under analysis have important policy implications.

A popular graphical test for detecting the presence of this form of publication bias is the funnel plot (Egger *et al.*, 1997). A funnel plot is a scatter diagram that compares the estimate of the effect size from each study, $\hat{\delta}_j$, (in the horizontal axis) against its precision (in the vertical axis), measured by the inverse of the estimate of the standard error of $\hat{\delta}_j$, $1/\hat{\omega}_j$. The logic behind funnel plots is the following. In the absence of publication bias and regardless of the magnitude of the "true" effect, estimates will vary randomly and symmetrically around it (recall the definition of unbiased estimator). On the other hand, the property of consistency establishes that estimates provided by studies with larger samples will be closer to the "true" effect size, while those provided by small-sample studies will be more spread out around the "true" effect size. Because small-sample studies with typically larger standard errors and hence less precision are at the bottom of the graph, in the absence of the publication bias the funnel plot should assume an inverted funnel-like pattern, symmetric around the "true" effect. However, if there is a bias in a certain direction, the graph will be asymmetric and overweighted (on the left or on the right side), especially in its bottom part. Thus, the key to identifying this form of bias is the funnel plot's asymmetry.

Figure 1 presents the funnel graph for the observations of our meta-sample. There does not seem to be evidence of the existence of publication bias in favour of a certain direction, as the graph appears to be fairly symmetric around the "true" effect size (which, in line with the fixed and random effects estimates of δ presented in subsection 4.1, is slightly negative).

Figure 1 - Funnel plot of the inequality-growth estimate



The funnel shape of the graph is unmistakable, and the symmetry seems to be clear. However, a closer inspection casts a modicum of doubt, as it reveals a thinner midsection for the left side of the funnel plot. Such close inspections of funnel plots reveal the fundamental weakness of using funnel graphs: visual inspections are inherently subjective and somewhat ambiguous (Stanley, 2008). This leads us to use a more objective tool to detect the presence of this form of publication bias. Egger *et al.* (1997) proposed a formal test for detecting asymmetry of the funnel plot - the Funnel Asymmetry Test (FAT). The FAT involves running a regression between a study's reported effect size and its estimated standard error:

$$\hat{\delta}_j = \gamma_0 + \gamma_1 \hat{\omega}_j + \varepsilon_j \tag{4}$$

In the absence of publication bias, $\hat{\delta}_j$ should vary randomly and symmetrically around the "true" value, γ_0, independently of $\hat{\omega}_j$. That is, parameter γ_1 in regression (4) should be equal to zero. When there is publication bias in a certain direction, studies with smaller samples and hence higher standard deviations tend to report an effect biased towards that direction. In this case, $\hat{\omega}_j$ should be statistically significant and parameter γ_1 significantly different from zero. Thus, the conventional t-test of the coefficient γ_1 ($H_0 : \gamma_1 = 0$ *vs.* $H_1 : \gamma_1 \neq 0$) is a test for publication bias, and the signal of its estimate indicates the direction of the bias.

However, equation (4) has a problem of heteroscedasticity. Given that its dependent variable is an estimated regression coefficient drawn from each original model in the primary literature, its estimated standard error, $\hat{\omega}_j$, is likely to vary with j.[7] As consequence, the estimated standard error of ε_j ($\hat{\omega}_j$, as well) is not constant. We can deal with this problem using the standard procedure of dividing regression (4) by $\hat{\omega}_j$, which yields:

$$t_j = \gamma_1 + \gamma_0 \frac{1}{\hat{\omega}_j} + \varepsilon_j^* \tag{5}$$

where $\varepsilon_j^* = \varepsilon_j / \hat{\omega}_j$ and $t_j = \hat{\delta}_j / \hat{\omega}_j$ is the conventional t-statistic associated to parameter δ, reported in the primary studies.[8] Thus, regression (5) should be used instead of (4), as in the former heteroscedasticity is likely to be eliminated. Note that because the intercept and the slope coefficients are now reversed, the FAT is now the t-test for the intercept γ_1.

Column (1) of Table 2 presents the results of the estimation of equation (5) by OLS for our meta-sample,[9] with standard-errors calculated using the Newey-West procedure. This procedure consistently estimates standard-errors in the presence of heteroscedasticity and/or non-specified autocorrelation between disturbances.[10] In our case, this technique is necessary, given the above mentioned presumable existence of correlation between observations drawn from the same study. We can see from Table 2 that the FAT confirms the previous interpretation of the funnel graph: the intercept of regression (5) is not significantly different from zero, which reveals little signs of publication bias in favour of a given direction.

[7]Recall that if the original model is estimated by, for example, OLS, the standard error of $\hat{\delta}_j$ is given by $\sigma_u^2 \left(Z'Z\right)^{-1}$, with Z standing for the matrix of the moderator variables in equation (2). Since the studies in the primary literature use different datasets, different sample sizes, and different independent variables, $(Z'Z)$ will differ from study to study and so will $\sigma_u^2 \left(Z'Z\right)^{-1}$.

[8]See in Column (4) of Table 1 the t-statistics associated to each observation of our meta-sample.

[9]Note that estimating (5) by OLS is equivalent to estimating (4) by WLS, with weights given by $1/\hat{\omega}_j$.

[10]From now on, we refer to these as heteroscedasticity-autocorrelation consistent standard errors. See Newey and West (1987) for a discussion on the calculation of a covariance matrix with such properties.

Table 2 - Results of the tests detecting for the presence of publication bias

Columns	(1)	(2)	(3)	(4)	(5)	(6)
Testing for:	**Direction Bias**	**Magnitude bias**	**Time patterns**	**Nationality patterns**	**Journal type patterns**	**Journal impact factor patterns**
Dependent variable	t_j	$\mid t_j \mid$	t_j	t_j	t_j	t_j
Constant	-0.6702 (-1.6388)	1.7714^{**} (7.3804)				
$\mathbf{1/\hat{\omega}_j}$	-0.0082 (-1.9267)	0.0059^{*} (2.1120)	-0.0503^{**} (-4.1164)	-0.0112^{**} (-3.6762)	-0.0185^{**} (-3.4971)	-0.0088 (-1.9157)
$\mathbf{time_j^*}$			0.0087^{**} (3.4620)			
$\mathbf{time_j^{2*}}$			-0.0004^{**} (-3.4497)			
$\mathbf{Europe_j^*}$				0.0033 (1.0467)		
$\mathbf{Others_j^*}$				0.0070 (1.8749)		
$\mathbf{Develop_j^*}$					0.0083 (1.2078)	
$\mathbf{Growth_j^*}$					0.0061 (0.4281)	
$\mathbf{Impact_j^*}$						-0.0017 (-0.0592)

Notes: Coefficients are estimated by OLS.
Moderator variable x_j^*corresponds to variable x_j divided by $\hat{\omega}$.
t-statistics in parenthesis and calculated from heteroscedasticity-autocorrelation consistent standard errors.
** and * denote statistical significance at the 1% and 5% level, respectively.

4.3.2 Bias in the magnitude of the results

Another important form of bias arises when studies are published more readily if they contain statistically significant results. In fact, because researchers, reviewers, and editors are usually predisposed to treat statistically significant results more favourably, these are more likely to be published. Studies that find relatively small and insignificant effects are much less likely to be published, as they may be thought to say little about the phenomenon in question. The problem for intelligent summary or review is that such bias leads to a truncated pool of published studies and makes empirical effects seem larger than they are.

This form of publication bias, also known as the "file drawer problem", has long been a major concern to meta-analysts. For nearly a half century, medical researchers and social scientists have acknowledged the seriousness of this problem (Sterling, 1959; Rosenthal, 1979; Begg and Berlin, 1988). At least since De Long and Lang (1992), economists have uncovered significant publication selection bias in many research areas (Card and Krueger, 1995; Ashanfelter *et al.*, 1999; Gorg and Strobl, 2001; Doucouliagos *et al.*, 2005; Docouliagos, 2005; Rose and Stanley, 2005; Stanley, 2008).

It should be noted that this bias need not arise because of the deliberate suppression of insignificant results, motivated by some urge to deceive. Authors may, for example, refrain from submitting statistically insignificant results on the expectation that they will have a lower probability of publication (Sutton *et al.*, 2000). Insignificant results may not be as interesting to readers and, given that journal space is a scarce resource, journals may prefer that insignificant results are not published, choosing instead to devote space to what are regarded more informative results (Doucouliagos *et al.*, 2005).

In general, studies that have smaller samples are at a distinct disadvantage of finding statistical significance. Indeed, because their standard errors are predictably larger, they will find it more difficult to produce high t-statistics (recall that $t = \hat{\delta}/\hat{\omega}$), whether or not a "true" effect exists. Hence, when there is a search for statistical significance, authors of small-sample studies may be tempted to manipulate their specifications (for example, by varying functional forms or changing the set of included covariates) in order to find required large estimates of the effect size. In turn, authors of studies with larger sample sizes and smaller standard errors will generally not need to search quite so hard, as lower values of the effect size are compatible with high t-statistics. Therefore, in the presence of publication bias towards statistical significance, one should expect a positive relationship between the magnitude of a study's estimate of the effect size and its standard error (Card and Krueger, 1995; Ashenfelter *et al.*, 1999; Gorg and Strobl, 2001; Stanley, 2005). This provides the basis for testing this form of bias, which involves estimating the regression:

$$| \hat{\delta}_j | = \gamma_0 + \gamma_1 \hat{\omega}_j + \varepsilon_j \tag{6}$$

This regression is very similar to (4), the only difference being that the dependent variable is the absolute value of $\hat{\delta}$. This is because now we are not interested in analyzing the direction of the bias, but the magnitude of $\hat{\delta}$, regardless of its signal. The test for this form of publication bias is thus $H_0 : \gamma_1 = 0$ (absence of bias) *vs.* $H_1 : \gamma_1 > 0$ (presence of bias). As previously, heteroscedasticity in (4) requires the estimation of:

$$| t_j | = \gamma_1 + \gamma_0 \frac{1}{\hat{\omega}_j} + \varepsilon_j^* \tag{7}$$

The estimation results of (7) for our meta-sample are presented in Column (2) of Table 2. The null hypothesis $H_0 : \gamma_1 = 0$ is rejected in favor of $H_1 : \gamma_1 > 0$, at 1% of significance. There is thus clear evidence of the existence of publication bias towards statistical significance; that is, the reported results of the effect of inequality on growth are likely to be overstated as a result of the desire to report and publish significant results.

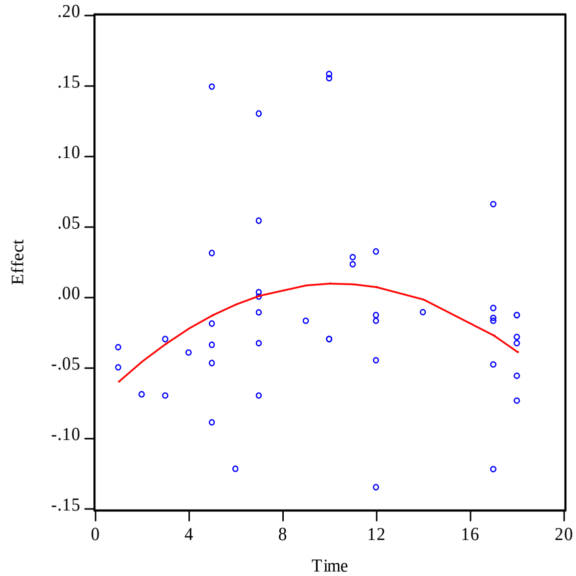
In this context, it is necessary to correct the estimates of δ provided by the primary studies. To do so, we follow Stanley (2005) procedure of, first, estimating the magnitude of each observation's bias (which is given by $\hat{\gamma}_1 \hat{\omega}_j$), and, second, shrinking each reported effect size towards 0 by $\hat{\gamma}_1 \hat{\omega}_j$. The corrected effect size of each observation of our meta-sample, $\hat{\delta}_j^{'}$, and the associated corrected t-statistic, $t_j^{'}$, are listed in Columns (5) and (6) of Table 1.

4.3.3 Publication patterns across time

Goldfarb (1995) argues that empirical economics research has a predictable time pattern of fashion and novelty. According to this pattern, initially researchers tend to report evidence that confirms a recently offered hypothesis. Then, after a sufficient passage of time for confirmations to accumulate (typically years), further confirmation is thought to contain little new information and will not be deemed to be interesting or publishable by reviewers and editors. Thus, contradictions become more likely to be published. Again, after sufficient time elapses, such empirical criticisms will become old-fashioned, engendering another reversal of publication preferences. This defines the "economics research cycle", which suggests that there are predictable cycles of fashion and novelty in empirical economics research.

Figure 2 reflects precisely the time pattern implicit in Goldfarb's conjecture. In the early-mid 1990s, when the inequality-growth empirical literature began to emerge, the tendency was for authors and journals to publish negative effects. This was possibly associated to the desire of supporting the flourishing theoretical literature, which at that time began to model the transmission channels between inequality and growth, predicting, in general, a negative relationship. With the passage of time, this tendency to publish negative effects was reversed, as studies reporting a positive effect became increasingly important since the beginning of the century. In recent years, several studies have attempted to conciliate the two perspectives, hence both positive and negative findings have been evenly produced and reported. Thus, there seems to be an "economics research cycle" in the estimation of the effect of inequality on growth, which is reflected in Figure 2 by a quadratic time pattern.

Figure 2 - Distribution of the reported effect sizes over time



In order to test formally Goldfarb's conjecture, we propose running a regression in which the reported effect size depends on a constant and on the variables $time_j$ and $time_j^2$, with $time_j$ defined as the year of publication of study j minus 1993.[11] The heteroscedasticity-corrected version of such regression is:

$$t_j = \gamma_0 \frac{1}{\hat{\omega}_j} + \gamma_1 time_j^* + \gamma_2 time_j^{2*} + \varepsilon_j^* \tag{8}$$

where $time_j^* = time_j/\hat{\omega}_j$, $time_j^{2*} = time_j^2/\hat{\omega}_j$ and $\varepsilon_j^* = \varepsilon_j/\hat{\omega}_j$. Column (3) of Table (2) shows that the estimate of γ_2 is negative and both $time_j$ and $time_j^2$ are significant at the 1% level, which confirms the statistical significance of the quadratic time trend reflected in Figure 2. Therefore, we can conclude that the inequality-growth empirical literature exhibits a time pattern "dictated" by Goldfarb's "economics research cycle", which constitutes another source of publication bias.

4.3.4 Publication patterns across nationalities

Several authors have mentioned nationality as a potentially important determinant of research findings. Neary *et al.* (2003) and Coupé (2003) suggest that there is a large difference in the quality of economists' output between Europe and US but this gap is shrinking. On a related issue, Stanley (2005) finds that there is a greater tendency to engage in publication selection bias among US studies. We

[11]Thus, $time_j$ assumes the value 1 for a study published in 1994, the value 2 for a study published in 1995, and so on.

address this issue by examining whether the reported effects of inequality on growth differ significantly according to the country of the institutions to which authors are affiliated.

To do so, taking into account that in our sample a considerable amount of the authors are affiliated to US and European institutions, we first divided it into three groups: Group 1, composed by those studies in which all the authors are affiliated to US institutions; Group 2, which comprises those studies in which at least one of the authors is affiliated to European institutions; Group 3, composed by the remaining studies. We then defined two dummy variables: variable $Europe_j$ assumes the value 1 if the study j belongs to Group 2, and variable $Others_j$ assumes the value 1 if the study j belongs to Group 3. Finally, we estimated by a OLS the following heteroscedasticity-corrected regression:

$$t_j = \gamma_0 \frac{1}{\hat{\omega}_j} + \gamma_1 Europe^*_j + \gamma_2 Others^*_j + \varepsilon^*_j \tag{9}$$

where $Europe^*_j = Europe_j/\hat{\omega}_j$, $Others^*_j = Others/\hat{\omega}_j$ and $\varepsilon^*_j = \varepsilon_j/\hat{\omega}_j$. As shown in Column (4) of Table 2 variables $Europe_j$ and $Others_j$ are not statistically significant at the 5% level, which means that the reported effect sizes of Group 2 and Group 3 are not significantly different from those of Group 1 (the reference category). Thus, there is no statistical evidence to support the idea that the reported effects of inequality on growth differ according to the country of the institutions to which authors are affiliated.

4.3.5 Publication patterns across journals

The last type of publication bias we investigate relates to the type of journals in which studies are published. Doucouliagos *et al.* (2005) suggest that journals from different fields may publish considerably different results. For example, while analyzing the results of the empirical literature on the union-productivity effects, they found that management journals publish primarily positive effects, while economics and industrial relation journals tend to publish both positive and negative effects. In order to test for the presence of this type of bias in the inequality-growth literature, we used the *Categorization of Journals in Economics and Management* done by the French Committee of Scientific Research and grouped the journals in our sample into three categories: general economics, development, and economic growth. We then proceeded as in subsection 4.3.4. General economics was defined as the reference category, while dummies $Develop_j$ and $Growth_j$ were created to indicate journals from the other two categories. A regression similar to (9) was estimated, with these two dummies as explanatory variables. The results, presented in Column (5) of Table 2, show that there are not significant differences in the reported results among these three types of journals.

We also investigated whether results might differ according to the ranking of the journals. Using the ISI Web of Science's 2010 impact factor as explanatory variable, we found no evidence of a relationship between the reported effect sizes and the ranking (see Column (6) of Table 2).

Therefore, we can conclude that there is no bias across journals in the empirical literature of the impact of inequality on growth.

4.4 Testing for the existence of an overall "true" effect size

In subsection 4.1 we calculated the fixed and the random effects estimates of the impact of inequality on growth and found that both are negative. In subsection 4.2 we came to the conclusion that there is not one but possibly several different "true" effect sizes, hence the random effects estimator is more appropriate. Subsection 4.3, in turn, showed that the reported estimates are likely to be distorted by some forms of publication bias. In this part of the paper we answer the following question: is the overall "true" effect size statistically different from zero after circumventing publication bias? That is, after removing the distortions caused by publication bias, does, on average, inequality really have a statistically significant impact on economic growth? We will answer this question through four different methods.

First, we can recalculate the random effects estimate using the corrected reported effect sizes, $\hat{\delta}_j^{\prime}$. The estimate remains negative (-0.0057), but is now clearly smaller in magnitude than that presented in subsection 4.1. Moreover, it is not statistically different from zero at the 5% level (p-value=0.0898). This is explained by the fact that the original reported effect sizes, $\hat{\delta}_j$, were overstated due to the tendency to report and publish significant results.

Second, taking into account the property of consistency of an estimator, we can look at the corrected reported effect sizes provided by studies with larger samples (or higher precision) so as to have an idea of the "true" effect size. Figure 2 suggests that there are four observations which detach from the remaining in what regards precision. The average reported effect size of these four observations is -0.00845, which is very close to zero.

Third, we can use equations (4) and (5) to identify the existence of an overall statistically significant "true" empirical effect, regardless of publication selection bias (in the direction and in the magnitude of the effect). As the sample size approaches infinity (or equivalently, $\hat{\omega}_j$ goes to zero), the expected value of the reported effect size approaches γ_0. Hence, γ_0 can be seen as a measure of the "true" effect size, after publication bias is filtered (Sutton *et al.*, 2000). The conventional t-test of the coefficient γ_0 ($H_0: \gamma_0 = 0$ *vs.* $H_1: \gamma_0 \neq 0$) is then a test for whether the overall "true" effect size is different from zero. Stanley (2005) calls this procedure the Precision Effect Testing (PET). Column (1) of Table 2 shows that γ_0 is not statistically different from zero, which corroborates the previous two conclusions.

Finally, we make use of another regression test, the Meta-Significance Testing (MST), specifically designed to identify the existence of a statistically significant "true" effect beyond publication bias. Stanley (2001) points out that if there is a real effect between two variables, then there should be a positive relationship between the natural logarithm of the absolute value of the t-statistic and the natural logarithm of the degrees of freedom in the regression:

$$ln \mid t_j \mid = \rho_0 + \rho_1 ln(df_j) + \epsilon_j \tag{10}$$

where df_j denotes the degrees of freedom from study j. Stanley (2005) shows that the slope coefficient in equation (10) offers information on the existence of a statistically significant "true" effect. If such effect exists, t-statistics are expected to rise as sample size rises and standard errors fall, hence $\rho_1 > 0$. More precisely, since we can expect a "double-log" relationship between a study's t-statistic and its degrees of freedom, ρ_1 should be close to 0.5. In the presence of publication bias (in the magnitude), this positive relationship is attenuated because studies with smaller samples tend to report larger t-statistics presumably in order to increase the prospects of publication. Stanley (2005) shows that, in this case, ρ_1 will decrease below 0.5 but remain greater than zero. Therefore, when reported effect sizes are overstated due to publication bias, (which is our case) the MST provides evidence of the existence of a significant overall "true" effect if $0 < \rho_1 < 0.5$; on the contrary, if such an effect does not exist, $\rho_1 \leq 0$. We estimated equation (10) by OLS and obtained a negative estimate for coefficient ρ_1 (-0.0276), which confirms the latter hypothesis.

In sum, the conclusion of this subsection is rather clear. All the four methods employed suggest that, after correcting for publication bias, the overall impact of inequality on growth is not statistically significant, meaning that, on average, the association between the two variables is not relevant.

4.5 Meta regression analysis - explaining the variation among reported estimates

At this point, based on the results obtained above, the conclusions can be summarized as follows.

First, by simply pooling the estimates in our sample and calculating their weighted average, we find an overall negative, albeit not economic meaningful, relationship between inequality and economic growth. Second, there are traces of publication bias in this empirical literature, as, on the one hand, authors and journals are more willing to report and publish statistically significant results, and, on the other hand, studies' results tend to follow a predictable cycle of fashion and novelty over time. Third, the overall effect of inequality on growth after correcting for these forms of bias is practically non-existent, meaning that, on average, the relationship between the two variables is not only economic meaningless but also statistically insignificant. Fourth, in spite of this fact, there may be not one but several "true" effects of inequality on growth, which are likely to differ in their nature and operate in opposing directions. This is suggested by the manifest heterogeneity in the reported results.

This subsection addresses precisely the issue of excess variation among effect sizes and the existence of a multiplicity of effects between inequality and growth. In particular, we will try to find what the sources of heterogeneity are and, based on this, draw some conclusions about the nature of such different effects. This will allow us to test objectively some of the conjectures made in Neves and Silva (2010) and, as a consequence, have a better understanding of the nature of the inequality-growth relationship and of its underlying transmission channels.

To do so, we make use of meta-regression analysis. As mentioned in Section 2, meta-regression is an adequate tool to model the heterogeneity in studies' findings, hence its increasing use in quantitative literature reviews in economics. Formally, it is represented by equation (1), in which the dependent variable is the estimate of the effect size reported by each study, and the independent variables include some characteristics of the studies.

We chose as moderator variables for our meta-regression those referring to characteristics that, according to our critical discussion in Neves and Silva (2010), are believed to explain the differences among effect sizes. Regarding the dependent variable, we used $\hat{\delta}'_j$ instead of $\hat{\delta}_j$, as the former filters publication bias from the reported effect sizes.[12] As in the previous estimations, the meta-regression was estimated by OLS, correcting for the presence of both heteroscedasticity (by dividing all the variables by $\hat{\omega}_j$) and auto-correlation (using the Newey-West procedure). Table 3 presents the estimation results, which are discussed below.

[12] It is important to note that this procedure only filters one form of bias, namely that resulting from giving preference to significant results. We correct for the existence of an "economic research cycle" by including $time_j$ and $time_j^2$ as additional explanatory variables.

Table 3 - Results of the meta-regression

Dependent variable: $t_j^{'}$	
Moderator variables	**Coefficient estimates**
$\mathbf{1/\hat{\omega}_j}$	-0.0295^{**} (-3.2287)
$\mathbf{time_j^*}$	0.0051^{**} (4.9112)
$\mathbf{time_j^{2*}}$	$-0.0002**$ (-3.8970)
$\mathbf{Panel_j^*}$	0.0053^{*} (2.5073)
$\mathbf{Regional.(1-Panel)_j^*}$	0.02157^{**} (4.1080)
$\mathbf{Developing_j^*}$	-0.0043^{**} (-2.8975)
$\mathbf{DevelopingOECD_j^*}$	-0.0072 (1.1550)
$\mathbf{Income_j^*}$	0.0127^{**} (3.0811)
$\mathbf{HQ.Income_j^*}$	-0.0077 (-1.7837)
$\mathbf{Expinc.Income_j^*}$	-0.0073^{*} (-2.3806)
$\mathbf{Specif_j^*}$	-0.0030 (-0.4590)
$\mathbf{Panel.Fixed_j^*}$	0.0484^{**} (14.7385)
$\mathbf{Panel.Random_j^*}$	0.0071 (0.5413)
N	45
$\mathbf{R^2}$	0.6835
$\mathbf{corrected-R^2}$	0.8798
$\mathbf{F-ratio}$	3.0146^{**}

Notes: Coefficients are estimated by OLS.
Moderator variable x_j^* corresponds to variable x_j divided by $\hat{\omega}$.
t-statistics in parenthesis and calculated from heteroscedasticity-autocorrelation consistent standard errors.
** and * denote statistical significance at the 1% and 5%, respectively.

One of the conclusions drawn in Neves and Silva (2010) was that the structure of the data used in the empirical studies is likely to influence the estimate of the effect size. In particular, cross-section studies typically report a negative and significant relationship between inequality and growth, while in panel studies the results are more diverse. We test this hypothesis by including in the meta-regression a dummy variable, labeled $Panel_j$, which assumes the value 1 if observation j is taken from a panel-data study and 0 if it is taken from a cross-section study. This variable is statistically significant at the 5% level and its coefficient estimate is positive, which confirms our inference. In Neves and Silva (2010), we advanced with two possible explanations for the divergences in the results of cross-section and panel studies. The first one relates to differences in the time horizon implicit in each type of studies: while the former examine the inequality-growth relationship in the long-run, the latter do so in the short-medium run, and the transmission channels between inequality and growth are likely to operate differently in both cases. The second explanation is that the effect of inequality on growth may differ substantially across countries and regions. Given that panel data, contrarily to cross-section data, controls for time-invariant unobservable country-specific characteristics, the existence of such specificities may make panel findings quite diverse.

If this second explanation is true, then the inclusion of regional dummies in cross-section studies should weaken the negative effect of inequality on growth. This was in fact one of the key ideas advanced in Neves and Silva (2010). To test for it, we include in the meta-regression another explanatory variable, $Regional_j$, which is equal to 1 if the primary study incorporates regional dummies in the base-regression, and equal to 0 otherwise. Given that this distinction applies only to cross-section studies, $Regional_j$ is multiplied by $(1 - Panel_j)$. The meta-regression shows that the coefficient associated to $Regional.(1 - Panel)_j$ has a positive estimate and is statistically different from zero, thereby confirming that the inclusion of regional dummies in cross-section studies weakens the impact of inequality on growth. It is worth noting that the effect of $Regional.(1 - Panel)_j$ is particularly strong, which is demonstrated by a low p-value (0.0003) and a high estimate of its coefficient (0.0216). This means that country and region specificities play a crucial role in explaining the heterogeneity found in the reported effect sizes.

We also investigate if the effect of inequality on growth differs according to the country's level of development. In Neves and Silva (2010), we suggested that this effect seems to be negative in developing countries (probably because of the importance of the credit market imperfections, sociopolitical instability and joint education/fertility channels) and insignificant or positive in developed ones (where the savings channel is likely to be equally important). The impact of using different samples of countries is examined by defining studies that include only OECD countries as the reference category. $Developing_j$ is a dummy that assumes the value 1 when the study includes only developing countries, and $DevelopingOECD_j$ is another dummy, equal to 1 when the study uses both types of countries. The results show that only $Developing_j$ is statistically significant. Its negative coefficient estimate suggests that in fact for less developed economies income inequality may hamper subsequent growth.

Another important inference made in Neves and Silva (2010) was that inequality in wealth (proxied by either land or human capital) seems to have a stronger negative impact on growth than inequality in income distribution. This idea is also clearly corroborated by the meta-regression, as dummy variable

$Income_j$ (which assumes the value 1 if the primary study uses income to measure inequality, and 0 if it uses wealth) is statistically significant, with a positive estimated coefficient. Again, transmission channels may be one of the explanations for this result, since wealth distribution is likely to be more relevant in those channels predicting a negative effect of inequality. Another possible explanation is related to the fact that the estimation result of the impact of income inequality on growth could be tainted by problems of measurement and comparability associated to data on income distribution. Let us further investigate this second line of argument.

As explained in Neves and Silva (2010), a common concern in the inequality-growth empirical literature is that data on income distribution are likely to be undermined by measurement error. In general, when a variable is badly measured, its coefficient is biased towards zero (*i.e.*, the so-called *attenuation* effect) resulting on a weaker impact on the dependent variable. In multivariate regression models the consequences are even more serious, as the error in one of the independent variables also biases the coefficients of the other variables, although in an unknown direction (Greene, 2000). Regarding income distribution data, its quality and quantity improved substantially with the introduction of the Deininger and Squire (DS) dataset in 1996, which compiles data based on three criteria of reliability.[13] More recent datasets, such as the Luxembourg Income Study (LIS) and the World Income Inequality database (WIDER) also comprise data that meet these criteria, and therefore are also considered high-quality datasets. We address the problem associated with the reliability of income data by dividing the studies in our meta-sample in two groups, one using high-quality income distribution data, drawn from the DS, LIS and WIDER datasets, and the other using data that does not satisfy the reliability criteria. HQ_j is a dummy equal to 1 if the observation j belongs to the first group, and equal to 0 if it belongs to the second group. Since we exclude from this analysis data on wealth distribution, HQ_j is multiplied in the meta-regression by $Income_j$. Variable $HQ.Income_j$ is not statistically significant, which confirms the conjecture advanced in Neves and Silva (2010) that using high quality income datasets does not seem to make a difference in the estimation of the effect of inequality on growth.

Another aspect associated to the measurement of income inequality has to do with the definition of income that is used. As argued by Knowles (2005) and Milanovic (2005) comparability of income inequality across countries is seriously hindered by the use of different concepts of income, namely gross income *vs.* expenditure. Moreover, mixing gross income and expenditure data is likely to introduce a

[13] See the three criteria for high-quality data defined by Deininger and Squire (1996) in Neves and Silva (2010).

bias in the results because expenditures tend to be more equally distributed than gross income. In order to achieve a higher level of cross-country comparability, several authors have followed Deininger and Squire's (1996) suggestion of transforming the original data by adding 6.6 points to expenditure-based Gini coefficients, thereby obtaining a proxy for gross income-based Gini coefficients when they are not available.[14] In order to check whether the use of different concepts of income affects the impact of inequality on growth, we defined a dummy variable, labeled $Expinc_j$, which assumes the value 1 if the study's dataset uses inequality based on both gross income and expenditure, and 0 if it uses inequality based on gross income only (measured with or without Deininger and Squire's transformation). The coefficient associated to $Expinc.Income_j$ is statistically different from zero and presents a negative estimate. This means that when only gross income-based inequality is considered, the reported impact of inequality on growth is higher, which is an expected result given that the distribution of gross income is more unequal than the distribution of expenditure.

In Neves and Silva (2010), we also raised the question of whether differences in the specification of the growth regression used in the primary studies could influence the reported results. In order to assess the impact of inequality on growth, all the studies in our sample estimated a regression of the form of equation (2), where the output growth rate is explained as a function of inequality and a set of other variables, Z_m, widely accepted in the literature as important determinants of growth. Several of these studies adopt the standard Perotti (1996) specification, which includes the initial GDP *per capita*, the level of investment, and the level of human capital as explanatory variables. In this case, dummy variable $Specif_j$ takes the value 1. When the growth regression does not assume a Perotti-type specification, it takes the value 0. Table 3 shows that this dummy is not statistical significant, lending support to the idea advanced by Neves and Silva (2010) that differences in the specification of the growth regression do not interfere in the estimation of the effect sizes.

Finally, we check whether heterogeneity of the reported effect sizes can also be explained by differences in the estimation techniques employed in the primary studies. Given that all cross-section studies use the same technique (OLS), this question will be investigated only for panel studies, where a wider variety of estimation techniques is employed. The standard methods of panel estimation are fixed effects and random effects. The fixed effects estimates are calculated from differences within each country across time; the random effects estimates are more efficient, since they incorporate information across individual countries as well as across periods. The major drawback with random effects is that

[14]This procedure is based on Deininger and Squire's (1996) finding that, for reliable data, expenditure-based measures yield Gini coefficients that are on average smaller by 6.6 points than gross income-based measures.

it is consistent only if the country-specific effects are uncorrelated with the other explanatory variables. $Fixed_j$ and $Random_j$ are dummy variables equal to one when the primary study uses fixed or random effects, respectively, and equal to zero otherwise. Thus, the reference category refers to other panel estimation techniques, such as GMM, which is the most appropriate estimator to deal with problems of endogeneity, caused, for example, by reversal causality from growth to inequality. These two dummy variables enter the meta-regression multiplied by variable $Panel_j$. In Table 3 $Panel.Fixed_j$ is statistically significant at the 1% level, showing that estimation techniques are important in explaining differences in the heterogeneity of effect sizes. In particular, panel estimations using fixed effects lead to a higher estimate of the impact of inequality on growth, which can be partially explained by the fact that inequality tends to be highly persistent over time (see Partridge, 2005).

In sum, from the estimation results of the meta-regression, we can conclude that most of the ideas advanced in Neves and Silva (2010) are corroborated by the empirical evidence. Moreover, the meta-regression does a very good job in explaining the heterogeneity of the reported effects of inequality on growth: the F-ratio shows that the model is statistically significant overall at the 1% level, and, more importantly, the corrected R-squared indicates that a big part (about 88%) of the variation in the effect sizes is explained by the meta-regression.[15]

The meta-regression also passes several diagnostic tests. Despite the expected presence of serial correlation (which is confirmed by the Breusch-Godfrey's LM test), White's test finds no trace of heteroscedasticity, Ramsey's generic misspecification test does not detect any evidence of omitted-variable or simultaneous equation bias, and Jarque-Bera's test does not reject the hypothesis that the disturbance terms have a normal distribution (see Table 4). The particular specification of the meta-regression can be further examined by testing the variance of the residuals. Recall that the dependent variable, the t-statistic reported in the primary studies, has a standard normal distribution under the null hypothesis of no effect. Then, if all systematic variation due to misspecification and differences in econometric models and methods of the original studies is adequately represented by the meta-regression, the remaining error variance should be equal to one. Stanley (2001) argues that, in this context, a test for the error variance being equal to one represents an additional way of assessing

[15]Note that the R-squared presented in Table 3 (equal to 68%), is an incorrect reflection of the meta-regression's ability to explain the variation in reported effects of inequality on growth, $\hat{\delta}_j$, because the dependent variable of the meta-regression is not $\hat{\delta}_j$, but $t_j' = \hat{\delta}_j'/\hat{\omega}_j$. After using the estimates of the coefficients β given in Table 3 to predict the effect sizes corrected for the presence of publication bias, $\tilde{\delta}_j'$, and after adding to these the respective magnitude of the bias, we obtained the predicted values of the effect size, $\tilde{\delta}_j$, for each observation. The comparison of these values with the reported effect sizes resulted in a corrected R-squared somewhat higher, 88%.

the specification of the model and its fitness to explain the heterogeneity of the effect sizes. Again, we confirm that our meta-regression performs well in what regards these aspects, as the null hypothesis that the residuals variance is equal to one is not rejected.

Table 4 - Battery of diagnostic tests to meta-regression

Test	Test-statistic	P-value
White for heteroscedasticity	$\chi^2_{(26)} = 24.8300$	0.5286
Breusch-Godfrey LM for serial correlation	$\chi^2_{(3)} = 14.1778$	0.0027
Ramsey RESET for model specification	$F_{(2,32)} = 0.0143$	0.9056
Jarque-Bera for normality of disturbances	$\chi^2_{(2)} = 0.2461$	0.8842
Variance of residuals equal to one	$\chi^2_{(44)} = 55.4223$	0.1159

Notes: Given the considerable number of explanatory variables in the meta regression, cross-terms were excluded in White's heteroscedasticity test.

Since the maximum number of observations collected from each primary study is three, Breusch-Godfrey's LM test was executed with three lags.

Following the standard procedure, Ramsey's RESET test was executed with one fitted term.

5 Concluding remarks

We conducted a quantitative analysis of the empirical literature on the effects of inequality on growth, using tools offered by meta-analysis. Such analysis is of major relevance as it allows systematizing the findings and drawing conclusions more objectively in a research field marked by clear divergences in the results and in the methodologies employed.

One important conclusion of this meta-analysis is that results are distorted by publication bias, in two ways. On the one hand, authors and editors are more prone to report and publish statistically significant effects, which makes the empirical effect of inequality on growth seem larger than it actually is. On the other hand, the results in this literature tend to follow a predictable pattern over time, according to which negative and positive effects are reported following an expected cycle of fashion and novelty.

We also found that, after correcting for these forms of bias, the overall impact of inequality on growth becomes insignificant, both statistically and economically, which means that on average the relationship between the two variables is weak. However, such an approach is rather superficial and may be completely misleading, as it does not capture the multiplicity of elements that characterize this relationship and render it rather complex. The manifest heterogeneity found in the reported effect sizes is a clear evidence that such complexity does exist.

Using meta-regression analysis, we investigated what the sources of this heterogeneity are, and found that most of the hypotheses advanced in Neves and Silva (2010) are corroborated. The estimate of the effect of inequality on growth is influenced by some characteristics of the empirical studies, such as the structure of the data, the sample coverage, and the type of distribution considered. In particular, it was confirmed that the impact of inequality on growth is negative and more pronounced in cross section studies, in less developed countries, and when inequality in wealth distribution is considered. On the contrary, when panel data is used, the sample is composed mostly by developed countries, regional dummies are added to the growth regression, and income distribution is used instead of wealth distribution the impact of inequality on growth becomes insignificant or even positive. In addition to these elements, the meta-analysis showed that differences in estimation techniques and in the definition of income are also relevant in explaining heterogeneity.

This analysis of the sources of heterogeneity is important not only *per se*, as it allows explaining quantitatively why the results in this research field differ so much, but also because it provides deeper insights about the way inequality influences growth. In fact, on the basis of the meta-regression results, we can say that there is not one but several underlying effects, which are different in their nature and operate through different channels. For example, inequality affects growth differently in developing and developed countries, meaning that the transmission mechanisms are not the same in both types of countries. Besides, the fact that panel data studies lead to more diverse and less conclusive results than cross section studies suggests that, on the one hand, the inequality growth relationship is influenced by country/regional specificities, and, on the other hand, it acts differently in the short and in the long-run. Also, inequality in wealth distribution has a stronger negative impact on growth that inequality in income distribution possibly because, as explored in Neves and Silva (2010), the transmission channels that are relevant in both types of distribution are not the same.

These insights provide important guidelines for both researchers and policy makers. Researchers should not have the pretension of finding a single, global pattern on the relationship between inequality and growth, because such pattern does not exist. Instead, they should focus on investigating this relationship from a specific perspective, and then define in accordance the appropriate datasets and methodologies. Similarly, policy makers should take into account that the level of inequality may have an important effect on growth, but this effect is not unequivocal and varies with countries. Understanding the different mechanisms that connect the two variables and the circumstances under which they operate is crucial for correct policy guidance in this area.

References

Alesina, A., and Rodrik, D. (1994) Distributive Politics and Economic Growth. *The Quarterly Journal of Economics* 109(2), pp. 465-490.

Ashenfelter, O., Harmon, C., and Oosterbeek, H. (1999) A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labour Economics* 6, pp. 453–470.

Banerjee, A., and Duflo, E. (2003) Inequality and Growth: What Can the Data Say? *Journal of Economic Growth* 8(3), pp. 267-99.

Barro, R. (2000) Inequality and Growth in a Panel of Countries. *Journal of Economic Growth* 5, pp. 5-32.

Begg, C., and Berlin, J. (1988) Publication Bias: A Problem in Interpreting Medical Data. *Journal of Royal Statistical Society* 151, pp. 419–463.

Bengoa, M., and Sanchez-Robles, B. (2005) Does equality reduce growth? Some empirical evidence. *Applied Economic Letters* 12(8), pp. 479-483.

Bergh, J., Button, K., Nijkamp, P., and Pepping, G. (1997) *Meta-Analysis in Environmental Economics*. Kluwer Academic Press, Amsterdam.

Bijmolt, T., and Pieters, R. (2001) Meta-analysis in Marketing when Studies Contain Multiple Measurements. *Marketing Letters* 12, pp. 157–169.

Bleaney, M., and Nishiyama, A. (2002) Explaining Growth: a Contest Between Models. *Journal of Economic Growth* 7. pp. 43-46.

Card, D., and Krueger, A. (1995) Time-series minimum-wage studies: a meta-analysis. *American Economic Review* 85, pp. 238–243.

Castelló, A. (2010) Inequality and Growth in Advanced Economies: an Empirical Investigation. *Journal of Economic Inequality* 8, pp. 293-321.

Castelló, A., and Doménech, R. (2002) Human Capital Inequality and Economic Growth: some new evidence. *The Economic Journal* 112, pp. 187-200.

Chambers, D., and Krause, A. (2010) Is the Relationship Between Inequality and Growth Affected by Physical and Human Capital Accumulation? *Journal of Economic Inequality* 8, pp. 153-172.

Clarke, G. (1995) More Evidence on Income Distribution and Growth. *Journal of Development Economics* 47, pp. 403-27.

Coupé, T. (2003) Revealed performances: worldwide rankings of economists and economics departments, 1990–2000. *Journal of the European Economic Association* 1, pp. 1309–1345.

David, L., and Hopkins, M. (2011) The Institutional Foundations of Inequality and Growth. *Journal of Development Studies* 47(7), pp. 977-997.

De la Croix, D., and Doepke, M. (2003) Inequality and Growth: Why Differential Fertility Matters. *The American Economic Review* 93(4), pp. 1091-1113.

De Long, J., and Lang, K. (1992) Are all economic hypotheses false? *Journal of Political Economy* 100, pp. 1257–72.

Deininger, K., and Olinto, P. (2000) Asset Distribution, Inequality, and Growth. World Bank Policy Research Paper no. 2375.

Deininger, K., and Squire, L. (1996) A new data set measuring income inequality. *World Bank Economic Review* 10(3), pp. 563-591.

Deininger, K., and Squire, L. (1998) New Ways of Looking at old Issues: Inequality and Growth. *Journal of Development Economics* 57, pp. 259-287.

Doucouliagos, C. (2005) Publication bias in the economic freedom and economic growth literature. *Journal of Economic Surveys* 19, pp. 367–388.

Doucouliagos, C., Laroche, P. and Stanley, T. (2005) Publication bias in union-productivity research. *Industrial Relations* 60, pp. 320–346.

Egger, M., Smith, G., Scheider, M. and Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 316, pp. 629–634.

Forbes, K. (2000) A Reassessment of the Relationship between Inequality and Growth. *The American Economic Review* 90, pp. 869-887.

Galor, O., and Zang, H. (1997) Fertility, income distribution and economic growth: Theory and cross-country evidence. *Japan and the World Economy* 9, pp. 197-229.

Glass, G. (1976) Primary, Secondary, and Meta-Analysis of Research. *The Educational Researcher* 10, pp. 3-8.

Goldfarb, R. (1995) The Economist-as-Audience Needs a Methodology of Plausible Inference. *Journal of Economic Methodology* 2, pp. 201–222.

Gorg, H., and Strobl, E. (2001) Multinational companies and productivity spillovers: a meta-analysis. *Economic Journal* 111, pp. 723–740.

Greene, W. (2000) *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall International.

Greenland S., and O'Rourke, K. (2008) In: *Modern Epidemiology* (Rothman K., Greenland S., Lash T., eds). Philadelphia: Lippincott Williams, pp. 652–682. Meta-analysis.

Hedges, L. (1982) Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics* 7, pp. 119-137.

Hedges, L., and Olkin, I. (1985) *Statistical Methods for Meta-Analysis*. New York: Academic Press.

Herzer, D., and Vollmer, S. (2012) Inequality and growth: evidence from panel cointegration. *Journal of Economic Inequality* 10(4), pp. 489-503.

Hunt, M. (1997) *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.

Hunter, J., and Schmidt, F. (1990) *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park: Sage Publications.

Jarrell, S., and Stanley, T. (1990) A meta-analysis of the union–nonunionwage gap. *Industrial and Labor Relations Review* 44, pp. 54–67.

Jarrell, S., and Stanley, T. (2004) Declining bias and gender wage discrimination? A meta-regression analysis. *Journal of Human Resources* 38, pp. 828–838.

Khalifa, S., and El Hag, S. (2010) Income disparities, economic growth, and development as a threshold. *Journal of Economic Development* 35(2), pp. 23-36.

Knowles, S. (2005) Inequality and Economic Growth: the Empirical Relationship Reconsidered in the Light of Comparable Data. *Journal of Development Studies* 41(1), pp. 135-139.

Li, H., and Zou, H. (1998) Income Inequality is not Harmful for Growth: Theory and Evidence. *Review of Development Economics* 2(3), pp. 318-334.

Light, R., and Pillemer, D. (1984) *Summing up - the science of reviewing research*. Cambridge: Harvard University Press.

Neary, J., Mirrlees, J., and Tirole, J. (2003) Evaluating economics research in Europe: an introduction. *Journal of the European Economic Association* 1, pp. 1239–1249.

Neves, P., and Silva, S. (2010) Inequality and Growth: Uncovering the main conclusions from the empirics. FEP Working Paper no. 381.

Newey, W., and West, K. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, pp. 703-708.

Partridge, M. (2005) Does income affect U.S. state economic growth? *Journal of Regional Science*, 45(2), pp. 363-394.

Perotti, R. (1996) Growth, Income Distribution, and Democracy: What the Data Say. *Journal of Economic Growth* 1, pp. 149-187.

Rose, A., and Stanley, T. (2005) A Meta-Analysis of the Effect on Common Currency on International Trade. *Journal of Economic Surveys* 19, pp. 347–365.

Rosenberger, R., and Loomis, J. (2000) A meta-analysis for benefit transfer: in-sample tests of an outdoor recreation database. *Water Resources Research* 36, pp. 1097–1107.

Rosenthal, R. (1978) Combining results of independent studies. *Psychological Bulletin* 85, pp. 185–193.

Rosenthal, R. (1979) The file drawer problem and tolerance for null results. *Psychological Bulletin* 86(3), pp. 638–641.

Rosenthal, R. (1984) *Meta-Analytic Procedures for Social Research*. Beverly Hills: Sage.

Sarkar., D. (2007) The role of human capital on economic growth revisited. *Applied Economics Letters* 14, pp. 419-423.

Shadish, W. (1982) A review and critique of controlled studies of the effectiveness of preventive child health care. *Health Policy Quarterly* 2, pp. 24-52.

Smith, V., and Kaoru, Y. (1990) What have we learned since Hotelling's letter? A meta-analysis. *Economics Letters* 32, pp. 267–272.

Stanley, T. (1998) New wine in old bottles: a meta-analysis of Ricardian equivalence. *Southern Economic Journal* 64, pp. 713–727.

Stanley, T. (2001) Wheat from chaff: meta-analysis as quantitative literature review. *Journal of Economic Perspectives* 15, pp. 131–150.

Stanley, T. (2004) Does unemployment hysteresis falsify the natural rate hypothesis? A meta-regression analysis. *Journal of Economic Surveys* 18, pp. 589–612.

Stanley, T. (2005) Beyond publication bias. *Journal of Economic Surveys* 19, pp. 309–345.

Stanley, T., Doucouliagos, C., and Jarrell, S. (2008) Meta-regression analysis as the socio-economics of economics research. *Journal of Economic Surveys* 3, pp. 54–67.

Stanley, T., and Jarrell, S. (1989) Meta-regression analysis: a quantitative method of literature surveys. *Journal of Economic Surveys* 3, pp. 54–67.

Sterling, T. (1959) Publication decisions and their possible effects on inferences drawn from tests of significance. *Journal of the American Statistical Association* 54, pp. 30–34.

Sutton, A., Duval, S., Tweedie, R., Abrams, K., and Jones, D. (2000) Empirical assessment of effect of publication bias on meta-analyses. *British Medical Journal* 320, pp. 1574–1577.

Sylwester, K. (2000) Income Inequality, Education Expenditures, and Growth. *Journal of Development Economics* 63, pp. 379-398.

Tanninen, H. (1999) Income inequality, government expenditures and growth. *Applied Economics* 31, pp.1109-1117.

Voitchovsky, S. (2005) Does the Profile of Economic Inequality Matter for Economic Growth? Distinguishing between the Effects of Inequality in Different Parts of the Income Distribution. *Journal of Economic Growth* 10, pp. 273-296.

Woo, J. (2011) Growth, income distribution, and fiscal policy volatility. *Journal of Development Economics* 96, pp. 289-313.