WPS7494

Policy Research Working Paper 7494

Weak Instruments in Growth Regressions

Implications for Recent Cross-Country Evidence on Inequality and Growth

Aart Kraay

WORLD BANK GROUP
Development Research Group
Macroeconomics and Growth Team
November 2015

Policy Research Working Paper 7494

Abstract

This paper revisits four recent cross-country empirical studies on the effects of inequality on growth. All four studies report strongly significant negative effects, using the popular system generalized method of moments estimator that is frequently used in cross-country growth empirics. This paper shows that the internal instruments relied on by this estimator in these inequality-and-growth regressions are weak, and that weak instrument-consistent confidence sets for the effect of inequality on growth include a wide range of positive and negative values. This suggests that strong conclusions about the effect of inequality on growth—in either direction—cannot be drawn from these studies. This paper also systematically explores a wide range of alternative sets of internal instruments, and finds that problems of weak instruments are pervasive across these alternatives. More generally, the paper illustrates the importance of documenting instrument strength, basing inferences on procedures that are robust to weak instruments, and considering alternative instrument sets when using the system generalized method of moments estimator for cross-country growth empirics.

This paper is a product of the Macroeconomics and Growth Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at akraay@worldbank.org.

Weak Instruments in Growth Regressions:

Implications for Recent Cross-Country Evidence on Inequality and Growth

Aart Kraay
World Bank Development Research Group

Keywords: Inequality, growth, weak instruments

JEL Codes: O4, D30, C36

____

1818 H Street NW, Washington, DC 20433, akraay@worldbank.org. I am grateful to the authors of the papers reviewed here for sharing their data and Stata codes and for answering numerous questions, and to Sam Bazzi, Frank Kleibergen, Peter Lanjouw, Christoph Lakner, David Roodman, Mark Schaffer, and Frank Windmeijer for helpful discussions. I am also grateful to David Roodman and Mark Schaffer for the features they have added to their Stata `xtabond2` and `weakiv` commands which greatly facilitate the implementation of weak instrument diagnostics and weak instrument-robust inference for dynamic panel estimators.

1. Introduction

The past few years have seen a resurgence of interest among academics and policy makers in the causes of inequality and its consequences for economic development. A central question in the debates around inequality concerns the effects of inequality on economic growth, with a growing consensus that inequality is detrimental to growth. In this paper, I revisit four recent studies that have contributed to this consensus with evidence of a negative and statistically significant relationship between inequality and subsequent growth: Castelló-Climent (2010); Halter, Oechslin and Zweimüller (2014); Ostry, Berg and Tsangarides (2014); and Dabla-Norris, Kochhar, Suphaphiphat, Ricka, and Tsounta (2015). The first two are published papers appearing in peer-reviewed journals in the fields of inequality and growth, respectively. The second two are IMF Staff Discussion Notes and as such have been widely discussed in policy circles and in the media.

Credibly estimating the effects of inequality (or any other variable) on growth is difficult. The fundamental empirical challenge is that any observed correlation between inequality and growth could reflect a causal effect running from inequality to growth, or it could reflect causation running in the opposite direction from growth to inequality, or it could simply reflect the influence of some third factor driving both inequality and growth. Since the policy implications of these three channels are very different, the recent empirical literature has devoted a great deal of effort to disentangling them. These identification challenges are not unique to the inequality-and-growth literature, and are widely recognized throughout the empirical literature on determinants of economic growth across countries.

Many papers in the empirical growth literature, including the four inequality-and-growth studies reviewed here, have relied on the system generalized method of moments (SYS-GMM) estimator as an econometric tool to address these identification challenges.[1] This estimator is essentially a system of two linear instrumental variables regressions, one relating growth rates to levels of explanatory variables, and one relating changes in growth rates to changes in explanatory variables. In order to isolate causal effects, this estimator relies on a large set of lagged levels and differences of right-hand-side variables as internal instrumental variables.[2]

[1] This estimator was developed in a series of papers including Holtz-Eakin, Newey and Rosen (1988), Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998), and was first applied in the cross-country growth literature by Caselli, Esquivel and Lefort (1996).
[2] The use of external instruments for inequality (as opposed to lags of inequality as internal instruments) is rare in the inequality and growth literature. A notable exception is Easterly (2007) who uses agricultural endowments

As in all instrumental variables regressions, the validity of statistical inferences about the parameters of interest -- in this case, the effect of inequality on growth -- depends crucially on the strength of the relationship between the instruments and the endogenous explanatory variables. In particular, when instruments are weakly correlated with the endogenous explanatory variables, it is well known that instrumental variable estimators can be seriously biased and have distributions in finite samples that are very different from their conventional asymptotic normal approximations.[3] This in turn means that the interpretation of point estimates and conclusions about their significance based on conventional t-statistics can be misleading when instruments are weak.

To address this problem, a large econometrics literature on weak instruments has developed tests to diagnose instrument strength, together with new methodologies that produce valid inferences even when instruments are weak. While the application of these tools is rapidly becoming standard practice in empirical papers that rely on single-equation instrumental variable regressions with external instruments, it remains rare in the very popular applications of the SYS-GMM estimator to cross-country growth regressions. A likely reason for this gap is the absence of "off-the-shelf" tools for assessing instrument strength and constructing weak instrument-consistent confidence sets in the context of the SYS-GMM estimator. Fortunately, the authors of the popular `xtabond2` and `weakiv` routines in Stata have recently added options that make it very easy to perform instrument strength diagnostics and generate weak instrument-consistent confidence sets in the context of SYS-GMM estimation. [4]

In this paper, I follow the approach of Bazzi and Clemens (2013) in applying weak instrument diagnostics to inequality and growth regressions.[5] Bazzi and Clemens (2013) unbundle the SYS-GMM estimator into its constituent "differenced" and "levels" equations and report weak instrument diagnostics and weak instrument-consistent inferences for these two equations separately, in a set of well-known empirical growth studies. Applying these tools to the benchmark specifications from the recent contributions to the inequality and growth literature considered here, I find pervasive evidence

(the relative share of land suited for wheat and sugarcane) as an instrument for inequality and documents a negative cross-sectional relationship between inequality and levels of development.

[3] An early reference to this observation is Nelson and Startz (1990). Since then a large literature on weak instrument-robust inference has developed. See Murray (2006) for a non-technical survey, and Mikusheva (2013) for a somewhat more technical but still accessible overview.

[4] See Roodman (2009) and Finlay, Magnusson and Schaffer (2013) for original documentation of these Stata commands, and the accompanying Stata help files for the descriptions of the most recent features. The replication code for this paper illustrates the convenience of the additional features in these Stata modules.

[5] See also Dollar and Kraay (2003) which reports weak instrument diagnostics and weak instrument-consistent confidence sets for differenced-GMM estimation of regressions of growth on trade and institutions.

of weak internal instruments in the SYS-GMM estimator used in these papers. Using weak instrument-robust approaches to inference, I find that the data used in these papers are consistent with a wide range of both positive and negative values for the causal effect of inequality on growth. This indicates that, once the poor performance of the internal instruments used in the SYS-GMM estimates of the benchmark specifications in these papers is properly accounted for, it no longer is possible to draw strong conclusions about either a negative or a positive effect of inequality on growth.

These findings in the benchmark specifications of these studies of course do not rule out the possibility that there may be other empirical specifications in which internal instruments are strong, or even if they are weak, that weak instrument-consistent confidence sets are small enough to be able to draw interesting conclusions about the growth effects of inequality. Rather, these findings illustrate the importance of systematically documenting instrument strength in all applications of the SYS-GMM estimator and relying on weak instrument robust approaches to inference where appropriate.

One key feature of the SYS-GMM estimator is that the underlying identifying assumptions imply that a large number of lagged levels and differences of explanatory variables are available as instruments. This in turn implies that users of SYS-GMM have a great deal of latitude to construct the instrument set from the multitudinous different possible combinations of these instruments, all of which are equally valid in terms of satisfying the relevant exclusion restrictions. This in turn raises the question of whether my conclusions about weak instruments hold not just for the instrument sets used in the benchmark specifications in these papers, but also across the many possible other instrument sets that could have been used instead. I take up this question in Section 4 of the paper, using the data set from Ostry, Berg, and Tsangarides (2014) as an illustration. Using this data set and a minimalist inequality-and-growth specification, I systematically document instrument strength across a large number of alternative combinations of internal instruments. I find that weak instrument problems appear to be pervasive across these alternative instrument sets, and that virtually none of them support strong conclusions about a significant effect of inequality on growth, either positive or negative. This stands in stark contrast with the SYS-GMM point estimates and t-statistics, which suggest a very consistent pattern of conventional significance across all instrument sets.

The weak instrument problems that I document in this paper arise in a specific set of influential recent contributions to the inequality and growth literature. Moving beyond this particular application, my findings underscore the conclusions of Bazzi and Clemens (2013) on the importance of verifying instrument strength, and if necessary also relying on techniques that are valid when instruments are

weak, in all applications of the SYS-GMM estimator to cross-country growth empirics. Beyond this, my analysis in the last section of the paper also emphasizes the importance of systematically verifying that conclusions are robust over a reasonably large set of alternative possible instrument sets. Paying more careful attention to issues of internal instrument strength will help to improve the credibility of empirical evidence based on the SYS-GMM estimator, not only in the inequality and growth literature, but in a much broader range of applications as well.

The rest of this paper proceeds as follows. In the next section, I summarize the findings of the four empirical studies reviewed in this paper. I also briefly review the SYS-GMM estimator and the use of weak instrument diagnostics and weak instrument robust approaches to inference in this setting. Section 3 contains my main empirical results, which show that weak instrument-consistent confidence sets for the estimated effect of inequality on growth in the benchmark specifications of these papers are very large and include a wide range of positive and negative values. In Section 4, I investigate how the performance of the internal instruments in the SYS-GMM estimator and the size of weak instrument-consistent confidence sets for the estimated effects of inequality on growth depend on some of the many possible choices for constructing the instrument set in the SYS-GMM estimator. Section 5 offers concluding remarks.

2. Background

2.1 Four Recent Inequality and Growth Studies

In this section I briefly summarize the main findings of the four studies considered here.[6] All four studies use the SYS-GMM estimator to identify the effects of inequality on growth, using cross-country panel data sets in which each observation is a five-year period for a given country. However, the studies differ in the specific hypotheses being tested, in their choice of inequality measures, and in their choice of control variables.

[6] Three other prominent earlier papers in this literature are similar to the ones considered here, in the sense that they also use lags of variables as internal instruments. Barro (2005) uses lags of inequality as instruments in a 3SLS regression but does not implement either the differenced or the levels equation that make up the SYS-GMM estimator. Forbes (2000) uses the differenced GMM estimator, and as such is likely to suffer from concerns about instrument strength similar to those documented here in more recent papers. Parenthetically, this paper is notable for being the only one among those discussed here to find a positive effect of inequality on growth. Interestingly, in all of the papers I replicate below, the differenced equation yields a positive (although not often significant) correlation between inequality and growth, consistent with difference-GMM estimates of Forbes (2000). Finally, Voitchovsky (2005) implements the full SYS-GMM estimator, and Bazzi and Clemens (2013) demonstrate that the internal instruments are weak in this paper, so I do not repeat this analysis here.

Ostry, Berg and Tsangarides (2014) (hereafter, OBT) emphasize the distinction between "gross" and "net" inequality (i.e. inequality before and after the redistributive effects of taxes and transfers are taken into account). They regress growth on initial income, a proxy for the net Gini, and a proxy for redistribution based on the difference between the net and the gross Gini. They rely on a data set described in Solt (2009) which extensively imputes the limited available cross-country data on net and gross income inequality to arrive at a very large and dense cross-country panel of annual observations on the net and gross Gini.[7] OBT find that higher net inequality is significantly associated with lower subsequent growth, while redistribution is not significantly correlated with growth. They conclude from this that redistributive policies that reduce the net Gini coefficient can increase growth.

Halter, Oechslin and Zweimüller (2014) (hereafter, HOZ) focus on the dynamic relationship between inequality and growth. They develop a theoretical model in which higher inequality is associated with faster growth in the short term, reflecting a more efficient allocation of assets across individuals. However, in the long term, an adverse political economy channel takes over, in which greater inequality leads to underprovision of public goods that subsequently reduces growth. To bring this theory to the data, they estimate a regression of growth on lagged inequality and twice-lagged inequality, taking their inequality data from the widely-used UNU-WIDER World Income Inequality Database. Consistent with their theory, they find that the coefficient on lagged inequality is positive (and sometimes significantly so), while the coefficient on twice lagged inequality is significantly negative. They also find that the sum of these two coefficients is negative, suggesting a negative impact of a permanent increase in inequality on long-run growth.

Dabla-Norris, Kochhar, Suphaphihpat, Ricka, and Tsounta (2015) (hereafter, DKSRT) provide a theoretical and empirical review of the causes and consequences of inequality, focusing on cross-country data. In the first part of their paper, they report new empirical evidence on the effects of inequality on subsequent growth. They first document a statistically-significant negative relationship between the net Gini and subsequent growth using an empirical specification and dataset that is very similar to OBT. They then use a smaller sample of countries for which income shares by quintile of the population are available in the UNU-WIDER World Income Inequality Database, and estimate a series of

[7] Since my focus in this paper is solely on the role of instrument strength, I do not delve into the various concerns about this dataset -- for details see Jenkins (2014). Briefly, the fundamental challenge with this imputed dataset is that actual data on pre- and post-tax inequality is very scarce outside OECD economies, and so for the developing economies that make up the majority of observations in OBT's data, the Solt (2009) data contains very little unimputed information about the extent to which developing countries redistribute through their tax and transfer systems.

regressions of growth on each of the five individual quintile shares, including them one at a time in their growth regressions. They find that higher bottom quintile shares are positively correlated with subsequent growth, while higher upper quintile shares are negatively correlated with subsequent growth. This latter finding is not very surprising given that top (bottom) quintile shares can be thought of as alternative summary measures of inequality that are almost perfectly positively (negatively) correlated with the Gini coefficient in their data.

Finally, Castelló-Climent (2010) (hereafter, CC) also documents the effects of inequality on growth in a large panel of countries. This paper considers inequality in the distribution of human capital, in addition to the income or consumption inequality measures that are studied in the other three papers reviewed here. Human capital inequality is constructed using the methodology proposed in Castelló and Domenech (2002), which relies on the Barro-Lee cross-country data on years of schooling by age groups of the population, and then calculates inequality in education across age groups as the measure of human capital inequality. Although this measure captures only human capital inequality across (and not within) age groups, an advantage of this measure is that it has much greater country and time coverage than measures of income and consumption inequality from the UNU-Wider World Income Inequality Database. Since the strongest results in this paper are for the human capital inequality measure, I focus on these here. CC finds a strongly significant negative relationship between human capital inequality and subsequent growth in a large panel of countries observed over five-year periods.

2.2 Estimating Growth Regressions Using the SYS-GMM Estimator

All of the studies discussed above estimate variants on the following standard cross-country growth regression:

$$y_{it} = \rho\, y_{it-1} + \beta' X_{it-1} + \mu_i + \lambda_t + \varepsilon_{it} \qquad (1)$$

where y_{it} is log real GDP per capita, X_{it-1} is a vector of explanatory variables, and the composite error term $\mu_i + \lambda_t + \varepsilon_{it}$ includes country and period effects as well as an idiosyncratic component. Note that subtracting lagged per capita income, y_{it-1}, from both sides of Equation (1) results in a regression of real per capita GDP growth on initial income and a set of control variables. In all four studies, the length of the time interval is five years. As noted above, the studies reviewed here differ in their choice of explanatory variables and sources of inequality data, but otherwise adhere to this basic setup. For

convenience, Table 1 summarizes the choice of explanatory variables included in the baseline specification of each study.

All four studies estimate this equation using the SYS-GMM estimator for dynamic panels with a short time dimension, developed by Holtz-Eakin, Newey and Rosen (1988), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998) (see also Roodman (2009) for an extensive discussion of this estimator and its implementation in Stata). Following the seminal application to cross-country growth empirics by Caselli, Esquivel and Lefort (1996), this methodology has been widely used in the empirical literature on the determinants of growth across countries. The estimator optimally combines inferences from GMM estimation of Equation (1) in levels and in differences, using appropriate lagged levels and differences of right-hand-side variables as internal instruments. Specifically, the SYS-GMM procedure estimates the "levels" Equation (1) jointly with a first-differenced version of Equation (1):

$$\Delta y_{it} = \rho\, \Delta y_{it-1} + \beta'\, \Delta X_{it-1} + \Delta\lambda_t + \Delta\varepsilon_{it} \tag{2}$$

where Δ is the first-difference operator, i.e. $\Delta x_{it} = x_{it} - x_{it-1}$ for any variable x. The rationale for differencing Equation (1) is that it eliminates the unobserved country-specific effect, μ_i, from the differenced equation This addresses the problem that μ_i is by construction correlated with lagged income on the right-hand-side, and moreover μ_i might also be correlated with the other explanatory variables of interest. However, differencing introduces a new challenge, which is that the differenced error term, $\Delta\varepsilon_{it}$, now is mechanically correlated with lagged growth, Δy_{it-1}, since the latter includes y_{it-1} which from Equation (1) is depends on ε_{it-1}, which in turn is included in the differenced error term. This problem is addressed with an appropriate choice of internal instruments, as discussed next.

SYS-GMM estimation of Equations (1) and (2) commonly relies on the following standard identifying assumptions. The first assumption is that the idiosyncratic component of the error term, ε_{it}, is uncorrelated over time. From Equation (1), y_{it} depends on current and lagged values of the error term, but not future values, so this implies that $E[y_{it-s}\Delta\varepsilon_{it}] = 0$ for s≥ 2, i.e. second and higher lags of log income per capita are valid instruments in the differenced equation. The second assumption is that $E[X_{it-1}\varepsilon_{it+s}] = 0$ for $s \geq 0$, i.e. the explanatory variables are predetermined with respect to the shock to growth. This implies that second and higher lags of the explanatory variables are valid instruments in the differenced equation. The third assumption, required to identify the levels equation, is that lagged

changes in explanatory variables are orthogonal to both the time-invariant and idiosyncratic component of the error term, i.e. $E[\Delta X_{it-s}(\mu_i + \varepsilon_{it})] = 0$ for $s \geq 1$. This implies that current and lagged differences of the right-hand side variables in the levels equation are valid instruments in the levels equation. Like all identifying assumptions, these three can be questioned on a variety of grounds. However, they are standard ones in the empirical literature that has used the SYS-GMM estimator, and I take them as given in what follows.

2.3 Weak Instruments

My focus in this paper is on the strength of the internal instruments implied by these identifying assumptions, i.e. do these instruments jointly have strong predictive power for the endogenous regressors? The answer to this question has important implications for the reliability of inferences about the estimated effects of inequality and redistribution on growth. Specifically, the literature on weak instruments has demonstrated that inferences based on conventional asymptotics applied to Wald statistics can be quite misleading when instruments are weak, and has proposed a variety of other approaches to inference that are robust to weak instruments. These techniques typically produce confidence sets for parameters of interest by "inverting" test statistics that are functions of the parameters of interest and have known asymptotic distributions regardless of instrument strength. In particular, confidence sets for parameters of interest can be derived from grid searches that identify regions of the parameter space where the test statistic is smaller than the appropriate critical value. These regions of the parameter space where the test statistic does not reject the null hypothesis at the α percent level form a $(100 - \alpha)$ percent confidence set. Since the validity of the asymptotic approximation does not depend on instrument strength for these tests, confidence sets constructed in this way are valid when instruments are strong and when they are weak.

Chernozhukov and Hansen (2008) provide a very intuitive interpretation of this approach. Consider an instrumental variables regression $y = X\beta + \varepsilon$ with a set of instruments Z. They observe that the Anderson and Rubin (1949) statistic, which can be used to construct weak instrument-consistent confidence sets, is equivalent to regressing $y - X\beta_0$ on the instruments Z and testing their joint significance. This is because under the joint null hypothesis that $\beta = \beta_0$ and the instruments are exogenous, i.e. $E[Z'\varepsilon] = 0$, the instruments should not be jointly significant in such a regression. Accordingly, a standard Wald test of the null hypothesis that the slope coefficients on the instruments are jointly zero is equivalent to a test of this joint null hypothesis. Moreover, since – roughly speaking -- such a test does not depend on the strength of the relationship between X and Z, it will be valid

regardless of instrument strength. Performing this test repeatedly over a grid of hypothesized-under-the-null parameter values β_0 traces out a confidence set for the parameters of interest that is asymptotically valid whether the instruments are strong or weak.

One drawback of confidence sets constructed in this way is that they are hard to visualize when there are more than two parameters of interest, as is the case in the papers considered here. For example, OBT include three endogenous regressors in their benchmark specification: lagged income, inequality, and redistribution. This implies that the joint confidence set for the coefficients on these three variables is a three-dimensional object that is difficult to represent graphically. To circumvent this problem, I follow the standard approach of projecting higher-dimensional confidence sets into lower dimensions. For example, in the case of OBT, consider a pair of hypothesized values for the slope coefficients on inequality and redistribution, which are the inequality variables of primary interest here. As long as there is at least one hypothesized value for the coefficient on initial income somewhere in the third dimension of the grid under consideration for which a joint test of all three values does not reject, this pair would be included in the two-dimensional joint confidence set for the coefficients on initial inequality and redistribution. Similarly, one can project into one-dimensional confidence sets for the individual parameters. For example, the one-dimensional confidence set for the slope coefficient on initial inequality consists of all hypothesized values for this parameter for which there is some pair of values for the other two slope coefficients such that the joint test of all three does not reject.

A drawback of such projection-based confidence sets is that they are conservative, in the sense that they include the true value of the parameter of interest more than 95 percent of the time if the joint test for each grid point in the parameter space is done at the 5 percent significance level. To ensure that my findings are not driven by this excess conservatism, I consider two alternative approaches as robustness checks. The first is to assume that the coefficient on initial income is strongly identified, following the methodology of Kleibergen (2004). This generates a weak instrument-consistent confidence set for only the weakly-identified parameters by evaluating the weak instrument-robust statistics over a smaller-dimensional grid of values for the only the weakly-identified parameters, at each point in the grid replacing the remaining strongly-identified parameter with an efficient estimate of it. An important caveat to this approach is that the assumption that the coefficient on initial income is strongly identified is unlikely to be literally true. The second alternative is to simply impose a value for the coefficient on initial income. Specifically, I assume that the coefficient on lagged income, ρ, is equal to its estimated value from the benchmark SYS-GMM estimates, $\hat{\rho}_{SYS}$, and then I regress $y_{it} -$

$\hat{\rho}_{SYS}\, y_{it-1}$ on $\beta' X_{it-1}$ using the same lagged levels and differences of the explanatory variables as instruments as in the SYS-GMM estimates. I then construct weak instrument-consistent confidence sets for the coefficients on the remaining inequality variables in the same way as before.

In the empirical results below, I report confidence sets based on three different statistics. The first is the Anderson-Rubin statistic, for which the intuition has already been discussed above. The literature has, however, noted that a shortcoming of this test is that it can lose power as the number of instruments becomes large relative to the number of endogenous regressors. This is particularly a concern in this application where the SYS-GMM estimator generates a large number of internal instruments. I therefore also present confidence sets constructed using two alternative techniques proposed to remedy this problem. Kleibergen (2005) proposes an orthogonal decomposition of the Anderson-Rubin statistic into (a) a K-statistic due to Kleibergen (2002) that tests only the null hypothesis that the parameters are equal to their hypothesized values, conditional on the orthogonality conditions being true, and (b) a J-statistic analogous to the standard test of the orthogonality conditions, but evaluated at the hypothesized parameter values. Kleibergen (2005) demonstrates that the power of the Anderson-Rubin test can be improved by separately testing these two null hypotheses at lower-than-5 percent significance levels at each point in the parameter space, and then constructing a confidence set consisting of those grid points that are not rejected by both tests. This yields 95 percent confidence sets with superior power properties to those based on the Anderson Rubin statistic alone. I also report confidence sets based on a conditional likelihood ratio statistic developed by Moreira (2003) for the i.i.d. case, and generalized by Kleibergen (2005), which tests only the null hypothesis about the slope coefficients of interest, conditional on the assumption that the orthogonality conditions are true. Moreira (2003) shows that this test also has better power properties than the one based on the Anderson-Rubin statistic when the number of instruments is large. As noted in the introduction, all of these confidence sets can readily be calculated using the `weakiv` module available in Stata, as described in Finlay, Magnusson and Schaffer (2013).

3. Empirical Results

3.1 Instrument Strength in the Baseline Specifications of OBT, HOZ, and CC

In this section, I document issues of instrument strength in three of the four papers described in Section 2.1.[8] I first replicate the benchmark empirical specifications of OBT, HOZ and CC, in the first columns of Tables 2, 3 and 4, respectively. OBT find a negative and significant estimate of the effect of net inequality on growth, and a small positive effect of redistribution on growth that is not significantly different from zero. They also reject the null hypothesis that these two slopes sum to zero, suggesting that redistributive policies that reduce the net Gini have a positive and statistically significant effect on growth. Their estimated negative growth effect of inequality is substantial. A one-standard deviation increase in the net Gini (i.e. a 0.1 point increase in the Gini on a scale from zero to one, corresponding roughly to the rise in inequality in China over the past 20 years) would lower cumulative growth over the following five years by 0.07 (i.e. 0.07/5=0.014, or 1.4 percent per year), corresponding to just over one-half a standard deviation of per capita GDP growth rates in the benchmark regression sample.[9]

In the case of HOZ, my replication is close to, but not exactly the same as the authors' results. I begin with the specification in Column 4 of Table 3 of their paper, which regresses log per capita GDP on its lag, as well as lagged and twice-lagged inequality, and a set of additional control variables. I make one correction and one change to this benchmark specification as estimated in the replication code provided by the authors. The correction involves including some instruments in the instrument set that were missing in the authors' original code.[10] The change is that I drop the additional control variables

[8] DKSRT kindly provided me with replication datasets and codes for the SYS-GMM inequality and growth regressions in their paper. However, I do not include this paper in my replication for two reasons. First, their specification and dataset are very similar to that of OBT, which I do replicate below, and so the same concerns about weak instruments are likely to apply in DKSRT as well. Second, the replication codes provided by DKSRT implemented the SYS-GMM estimator in a way that is inconsistent with the internal logic of this estimator. Specifically, the published regressions in DKSRT are implemented using the `xtdpdsys` command in Stata. In the replication code they provided to me, they used the following syntax: `xtdpdsys gdpgrowth log_gdp2005pc gini`. This formulation fails to recognize that the syntax of `xtdpdsys` treats variables listed immediately after the dependent variable as exogenous covariates, while predetermined regressors are listed subsequently using the `pre()` option. This implies that lagged income (`log_gdp2005pc`) and inequality (`gini`) are treated as strictly exogenous in the way that DKSRT used `xtdpdsys`. In the case of lagged income, this is inconsistent with the internal logic of the SYS-GMM estimator. In the case of inequality, this assumes away any potential endogeneity problems for inequality which motivated the use of the SYS-GMM estimator in the first place. This in turn implies that assessing internal instrument strength in their benchmark specification is not particularly meaningful, and so I do not pursue this further here.

[9] Note that OBT have average annual growth over a five-year period as the dependent variable in their specification. This means that the coefficients on inequality and redistribution in Table 2 are 5 times as large as the coefficients reported in OBT, which is just an inconsequential rescaling.

[10] HOZ implement the SYS-GMM estimator using `xtabond2`. The syntax of `xtabond2` requires that all exogenous regressors be included in the variable list following `xtabond2`, and also in the `ivstyle()` instrument list since they serve as their own instruments (see Roodman (2009), page 124). However, HOZ did not do this in their replication code for two sets of variables: (i) the period dummies, which I include in the `ivstyle()` instrument list for both the levels and differenced equations as required by `xtabond2` syntax, and

included in their benchmark specification, in order to focus more directly on the strength of the internal instruments for inequality. My version of HOZ's benchmark specification is reported in the first column of Table 3. Consistent with their findings, lagged inequality is positively but insignificantly correlated with growth, while twice-lagged inequality is negatively and strongly significantly correlated with growth. The main difference with HOZ's benchmark specification is that the estimated effect of twice-lagged inequality on growth is slightly more negative (-0.67 in my version, as opposed to -0.55 in theirs). The estimated long-run effect (the sum of the two coefficients) is negative and equal to -0.51. The magnitude of this estimated adverse growth effect of higher inequality is again substantial: a 0.1 point increase in the Gini would lower cumulative growth over the following five years by five percentage points, or about one percentage point per year.

The first column of Table 4 contains my replication of the benchmark specification in CC, which I take to be Column 1 in Table 1 of that paper. CC regresses growth on lagged income, lagged human capital inequality, and a set of control variables. Once again, I drop the control variables in order to focus directly on the strength of internal instruments for inequality. This change has only small effects on the estimated magnitude of the negative growth effect of human capital, which remains substantial: a one-standard deviation increase in human capital inequality of 0.24 Gini points in their sample lowers cumulative growth over the following five years by 5.7 percent, or about 1.1 percentage points per year. However, the significance of the estimated effect is lower than in CC's reported results, falling just short of significance at the 10 percent level.

At the bottom of Tables 2, 3 and 4, I detail the internal instruments used in the SYS-GMM estimator. In the differenced equation, OBT use the second through fifth lags of log per capita GDP, together with second and third lags of inequality and redistribution, as instruments. In the levels equation, OBT use the first and second lagged differences of inequality and redistribution as instruments. In the case of HOZ, the instruments in the differenced equation consist of twice-lagged levels of income, inequality, and the three control variables, while in the levels equation they use the lagged difference of income, inequality, and the control variables. For CC, the instrument set consists of

(ii) ΔG_{it-2} in the differenced equation, which serves as its own instrument under HOZ's maintained assumption that lagged inequality is predetermined, i.e. $E[G_{it-1}\varepsilon_{it+s}] = 0, s > 0$. This correction substantially strengthens identification relative to HOZ's benchmark specification, by introducing missing instruments that by construction are perfectly correlated with explanatory variables. However, it has only minimal effects on the point estimates in the SYS-GMM estimator.

the second- to fourth- lagged levels of right-hand-side variables in the differenced equations, and first-lagged differences in the levels equation.

In the second and third columns of Table 2, 3 and 4, I unbundle the SYS-GMM estimator into the corresponding differenced and levels equations, and report 2SLS estimates of each, using the same internal instruments for each equation as those used in the SYS-GMM estimator. In all three papers, the within-country variation exploited by the differenced equation suggests a positive relationship between inequality and growth, although this effect is statistically significant only in HOZ. On the other hand, the findings from the SYS-GMM estimator of a negative effect of inequality and growth appear to be driven primarily by the cross-country variation captured in the levels equation. For all three papers, the levels equation gives estimates of a negative effect of inequality on growth that are similar to, or slightly larger than, the SYS-GMM estimates. Accordingly, in much of the discussion below I will focus on instrument strength in the levels equation as it provides the best evidence in support of the main conclusions in these three papers.

I now turn to the question of instrument strength in these specification.[11] I first test the null hypothesis of underidentification using the Kleibergen and Paap (2006) rk-LM statistic. Specifically, the null hypothesis is that the matrix of coefficients from the first-stage regressions is not full rank, signaling a complete failure of identification. This is a fairly undemanding null hypothesis that, even if rejected, may still be consistent with the instruments having very weak explanatory power for the endogenous variables. In the case of OBT, the p-value for the null hypothesis of underidentification is 0.18 in the differenced equation, and 0.17 in the levels equation. Similarly, the null of underidentification is not rejected in both the differenced and levels equations for HOZ. Only in the case of CC is the null of underidentification rejected at the 5 percent level. The failure to reject even this undemanding null hypothesis of underidentification at conventional significance levels in these first two studies is a first indication of weak instrument problems.

I next report the Cragg-Donald statistic, and the corresponding critical values for the Stock and Yogo (2005) test of weak instruments. The critical values are for a test of the null hypothesis that the instruments are weak, where “weak” is defined in terms of a maximal bias of the 2SLS estimates relative

[11] As noted earlier, I take as given the identifying assumptions that generate the internal instruments in the SYS-GMM estimator. However, in the case of OBT it is worth noting that the Hansen test of overidentifying restrictions casts some doubt on these identifying assumptions, in the case of the levels equation which drives their results – the p-value for the null hypothesis that the instruments are orthogonal to the error term in the levels equation is just 0.036.

to the OLS estimates. I use the tabulated critical values for the least demanding version of this hypothesis, which allows for a 30 percent maximal relative bias.[12] The null hypothesis of weak instruments is rejected if the test statistic exceeds the corresponding critical value. For all three studies, and in both the differenced and the levels equation, I find that the null hypothesis of weak instruments cannot be rejected, even at the least demanding level of 30% bias.

One drawback of the Stock-Yogo approach is that the weak instrument set in the null hypotheses is defined in terms of a weighted average of the relative biases in all of the estimated coefficients. This may disguise the possibility that the coefficients on some variables are more strongly identified than others. To remedy this problem, I report conditional F-statistics for each first-stage regression proposed by Sanderson and Windmeijer (2015), that can be compared with the Stock-Yogo critical values coefficient-by-coefficient to test the null hypothesis of a weak instruments, where "weak" is defined in terms of maximal size distortions for a conventional Wald test of a null hypothesis involving that parameter only. In all but one case in Tables 2-4, the Sanderson-Windmeijer test fails to reject the null of weak instruments, even at the least demanding tabulated critical values corresponding to a maximal 25 percent size distortion. The only exception is in the levels equation of HOZ, where the null of weak instruments is rejected for the coefficient on lagged inequality (but not twice-lagged inequality).

A shared shortcoming of both the Stock-Yogo and Sanderson-Windmeijer tests for weak instruments is that they both are valid only under the assumption of i.i.d. errors. In contrast, the SYS-GMM estimator and the 2SLS estimates of the levels and differenced equations allow for the error terms to be heteroskedastic and correlated with each other within countries. For this reason, these two tests for weak instruments should only be interpreted as suggestive evidence of weak identification in this particular setting. Rather than pre-test for weak instruments, a more direct approach is to base inferences on statistics that are robust to the presence of weak instruments. In Tables 2-4, I report confidence sets for the coefficients on the main inequality variables of interest: lagged inequality and redistribution, in the case of OBT, lagged and twice-lagged inequality in the case of HOZ, and lagged human capital inequality in the case of CC. These confidence sets are based on the AR statistic, the KJ statistics, and the CLR statistics. As discussed in the previous section, these one-dimensional confidence sets are obtained by projecting the multi-dimensional confidence set into each axis, and therefore are

[12] Stock and Yogo (2005) tabulate these critical values only for the case of 1, 2, and 3 endogenous regressors, and for up to 100 instruments. The critical values decline slightly with the number of regressors, and I linearly extrapolate them to the case of 5 (HOZ) and 6 (CC) endogenous regressors.

more conservative than their stated 95 percent sizes. In addition, in Figures 1 and 2, I graph the corresponding two-dimensional confidence sets for the coefficients on the two inequality variables of interest in OBT and HOZ, projecting out only the dimension corresponding to the coefficient on lagged income. The corresponding one-dimensional confidence intervals in Tables 2-4 are obtained by projecting these sets onto the horizontal and vertical axes. I calculate the confidence sets over a grid of values for the coefficients on inequality ranging from -2 to 2, and for the coefficient on lagged income ranging from 0.75 to 0.99.

These weak instrument-consistent confidence sets for the coefficients on the inequality variables are large. In the case of OBT, the 95 percent weak instrument-consistent confidence set for the effect of inequality on growth contains the entire grid of values from -2 to 2. This holds for both the levels and differenced equations, and for all three test statistics. This implies that weak-instrument robust inference cannot rule out *any* value within this wide range of hypothesized effects of inequality on growth. To put the size of this range in perspective, OBT's point estimate of -0.71 implies that a 10 point increase in the Gini coefficient would reduce growth by 1.4 percent per year. In contrast, the weak instrument-consistent confidence set includes values of the growth effect of inequality of a 10 point increase in the Gini coefficient ranging from -4 percent per year to 4 percent per year (i.e. $\beta = \pm 2 \times 0.1 \div 5 = \pm 4$ percent per year). The confidence sets for the effects of redistribution on growth are also large, and consist of the entire grid from -2 to 2.

These patterns are also quite clear from the two-dimensional confidence regions for the growth effects of inequality (horizontal axis) and redistribution (vertical axis) displayed in Figure 1, for the levels equation. In this figure, the dark-shaded region represents the combinations of parameter values that are not rejected. While their shapes differ somewhat, the AR, KJ, and CLR-based confidence all include a wide range of positive and negative values for the effects of inequality and redistribution on growth. Crucially, these confidence sets are quite different from those based on the conventional Wald statistic, shown for reference in the top-left panel. This conventional confidence set is a small elliptical region covering only negative values for the growth effect of inequality, and a roughly even mix of positive and negative values for the effects of redistribution on growth. The difference between the Wald-based confidence sets and the three weak instrument-consistent confidence sets clearly show that accounting for low instrument strength can lead to very different inferences about the key parameters of interest.

A similar pattern emerges for HOZ. For the levels equation, the weak instrument-consistent projection-based confidence sets in Table 3 cover the entire grid of parameter values for the estimated

effect of twice-lagged inequality on growth, using the AR, KJ and CLR statistics. In the differenced equation, the confidence sets are somewhat more compact, but for all three statistics still include both positive and negative values for the estimated growth effect of lagged inequality. Turning to Figure 2, the two-dimensional weak instrument-consistent confidence sets for the coefficients on lagged and twice-lagged inequality are also quite large and irregular, and typically contain a wide range of positive and negative values for the estimated effects of both variables on growth.

In the case of CC, the weak instrument-consistent confidence sets for the effect of inequality on growth in the differenced equation are large and contain a wide range of positive and negative values. However, in the levels equation the confidence sets are quite small and contain only negative values in the case of the K-J and CLR statistics. This apparently significant estimated negative effect of inequality on growth is at first glance surprising, given that the weak instrument diagnostics discussed previously clearly suggested weak instrument problems. Upon closer inspection, it turns out that these small confidence sets appear to be an artifact of the fairly narrow range of grid points values for the coefficient on initial income. If this range is expanded from [0.75, 0.99] to [0.5, 1.5], the confidence set for the coefficient on inequality becomes much larger, and consists of values in the range [-2 ,-0.22] U [0.06, 0.26] U [0.46, 2], i.e. nearly the entire grid for this coefficient. This is because the data also cannot reject a wide range of values corresponding to a positive effect of inequality on growth as well as a coefficient on lagged income that is greater than one.

One drawback of both the one- and two-dimensional confidence sets discussed above is that they are all conservative, because they are constructed by projecting out the dimension corresponding to the coefficient on lagged income. As a robustness check, in Table 5 and Figure 3, I consider the two alternative methods for dealing with this coefficient discussed in the previous section. Figure 3 reports true 95 percent two-dimensional weak instrument-consistent confidence sets for the coefficients on the two inequality variables of interest in the levels equations of OBT and HOZ. In the first and third columns, I assume that the coefficient on initial income is strongly identified, while in the second and fourth columns, I impose a value for this coefficient equal to its point estimate from the SYS-GMM estimator. While the precise shapes of the confidence sets are different across these different choices, the main conclusion that they are large and contain a wide range of positive and negative values for the estimated effect of inequality on growth remains. Similarly, in Table 5, I report weak instrument-consistent one-dimensional confidence sets, for these two alternative treatments of the coefficient on

initial income. Again, I find that weak instrument-consistent confidence sets for the main inequality variables of interest are large, and always include zero.

In summary, all three papers replicated in this section report estimated effects of inequality on growth that are significantly negative, based on conventional t-statistics that rely on standard asymptotic approximations. However, the internal instruments used by the SYS-GMM estimator appear to be weak, undermining the validity of these asymptotic approximations and the inferences to which they lead. When inferences are based on weak instrument-robust statistics, it no longer is possible to rule out (at conventional significance levels) a wide range of both negative *and* positive effects of inequality on growth. This suggests that strong conclusions about the growth effects of inequality in one direction or the other cannot be drawn from the benchmark specifications in these three studies.

4. Strength of Alternative Internal Instrument Sets in the SYS-GMM Estimator

In the previous section, I examined instrument strength in the benchmark specifications of the papers under consideration and found that the strength of the internal instruments was a major concern. However, there are many options for specifying different combinations of lags of variables as internal instruments in the SYS-GMM estimator, and it is possible that identification could be stronger with alternative internal instrument sets. I systematically explore this possibility in this section of the paper, by examining the robustness of the conclusions of the previous section to various choices in defining the instrument sets. To keep things as simple as possible, I focus on the most parsimonious specification of an inequality-and-growth regression: a regression of log per capita GDP on lagged log per capita GDP and the lagged Gini coefficient, as well as a set of period dummies i.e. Equation (1) with $\beta' X_{it-1} = \beta G_{it-1}$. Even in this stripped-down specification, there are multitudinous possibilities for specifying differing sets of internal instruments. For example, in the differenced equation, any combination of twice or higher lagged values of the right-hand side variables can be used as instruments, while in the levels equation any combination of first or higher lags of differences of variables are available as instruments. Since in a large cross-country panel such as that used by OBT there are as many as 8 lags available, this implies that there are in principle 32 different instruments to choose from (8 lags x 2 variables x 2 equations). Since in principle all combinations of these instruments are valid instrument sets, this in turn implies 2^{32} or over 4 billion possible instrument sets. The possibilities proliferate further upon realizing that each lag of a variable can appear as a single variable in the instrument set (referred to as "collapsed" form in the SYS-GMM estimator), or as a separate variable for each time period after that lag (referred to as "GMM-style" in the SYS-GMM estimator).

To keep the problem of assessing instrument strength over this proliferation of possible instrument sets manageable, I consider only a subset of these possibilities. In the differenced equation, I consider instrument sets consisting of the second to $(k+1)^{th}$ lags of income and the Gini, for a total of k^2 possible combinations of instruments. Similarly, in the levels equation, I consider instrument sets including the first to k^{th} lagged differences of income and the Gini, for a total of k^2 combinations of instruments for this equation as well. In addition, for each of these $k^2 \times k^2 = k^4$ combinations of instruments for the levels and differenced equations together, I consider both the "collapsed" and non-"collapsed" versions of the instruments, for a total of $2k^4$ possible instrument sets. In the results below, I set $k = 5$ so there are 1250 different possible specifications corresponding to 1250 different instrument sets. These break down into 50 distinct instrument sets for the differenced equation, and 50 distinct instrument sets for the levels equation.

I implement this robustness analysis using the OBT data set as an illustration. As noted above, I consider the minimal specification of a regression of growth on lagged per capita GDP and lagged inequality, and so I drop the redistribution variable included in their main regressions. Since this variable does not enter significantly in their benchmark specification, dropping it has little effect on their main finding of a negative and significant effect of net inequality on growth. For each instrument set, I implement the SYS-GMM estimator and retrieve the estimated coefficient on inequality, together with its standard error. Table 6 summarizes the distribution of point estimates of the growth effects of inequality, together with the absolute value of their conventional t-statistics. Consistent with the benchmark specification in OBT, the estimated effect of inequality on growth from the SYS-GMM estimator, reported in the top panel of Table 6, is significantly negative in nearly all specifications. The SYS-GMM point estimates range from -1.4 (at the 10^{th} percentile) to 0.43 (at the 90^{th} percentile), with a median value of -0.8 which is very close to the OBT benchmark estimates of around -0.7.[13] The table also distinguishes between models with "GMM-style" and "collapsed" instrument sets, with slightly more negative effects when the instrument set is collapsed. The distribution of t-statistics shows that in

[13] It is worth noting that the OBT benchmark specification reported in Table 2 does not appear among the 1250 models summarized in Table 5. This is due to three different reasons: (a) as noted previously I do not include the lagged redistribution variable that is in OBT's benchmark specification, (b) OBT do not include the lagged change in log GDP per capita in the instrument set for the levels equation, whereas in my robustness analysis here I always include at least one lag, and (c) I include the first and higher lags of the change in inequality in the instrument set for the levels equation whereas OBT start with the second lag.

nearly all specifications, the SYS-GMM estimate of the effect of inequality on growth is negative and statistically significant.

The second and third panels of Table 6 report the distribution of results across the $2k^2 = 50$ possible instrument sets for the differenced and levels equations separately. I first summarize the distribution of the estimated growth effect of inequality from both equations, as well as the distribution of the t-statistics. Consistent with the findings from the previous section, the estimated growth effects of inequality are overwhelmingly negative and with t-statistics greater than two in absolute value in the levels equation. In contrast, the estimates tend to be positive and not significantly different from zero in the differenced equation. This pattern is very consistent with the benchmark specification in OBT, and also in the other papers considered here.

The key questions of interest for this paper concern the strength of the internal instruments and the implications for inferences about the growth effect of inequality. As in the previous section, I focus on the levels equation, since this is where the evidence for a negative growth effect of inequality appears to be strongest. Consider first tests of underidentification. The null hypothesis of underidentification is rejected for just over half of the 50 models. However, there is a stark difference between models with a small number of instruments (i.e. those in which the instrument set is collapsed) and with a large number of instruments (i.e. in which the instruments enter "GMM-style" with a separate instrument for each time period). The null of underidentification is rejected for 24 out of 25 models in the former category, but only for four out of 25 models in the latter category. This is a first indication that problems of weak instruments in this setting appear to be much more pronounced when the number of instruments is large.

I turn next to tests of weak instruments. In only four of the 50 models considered here does the Stock Yogo test reject the null of weak instruments at the least demanding critical value of 30 percent maximal 2SLS bias relative to OLS. Similarly, the Sanderson and Windmeijer test for weak instruments based on maximal size distortion for the coefficient on lagged inequality rejects the null of weak instruments for only 11 out of the 50 possible models for the levels equation. These weak instrument diagnostics suggest that the problem of weak instruments is pervasive in most of the alternative instrument sets considered here. Consistent with the findings on underidentification, it is also worth noting that the problem of weak instruments seems closely related to the number of instruments. The relatively few specifications in which the weak instrument diagnostic tests suggest that instrument strength is not a problem all feature a small number of internal instruments. This is because these

models tend to have a collapsed instrument set and a small number of lags included in the instruments, both of which drastically reduce the number of instruments.

It would be tempting to select a preferred model with the strongest internal instruments based on these criteria and to base inferences on this one pre-selected model. However, a pretesting procedure for selecting instruments based on the strength of the first-stage regression necessarily leads to violations of the exclusion restriction, because the selection rule for the instruments reflects their correlation with the endogenous variables. Rather than follow such a data-based rule for selecting the instrument set, I simply report the weak instrument-consistent confidence sets for the effect of inequality on growth from all 50 versions of the levels equation. These are summarized graphically in Figure 4, which consists of four panels corresponding to confidence sets based on the AR, KJ, and CLR tests, as well as conventional confidence sets based on a standard Wald test, i.e. the 2SLS point estimate +/- 1.96 times the 2SLS standard error. The horizontal axis in each panel contains the hypothesized values for β, on a grid from -2 to 2. Each horizontal line represents the values of β included in the confidence set for a given model. Models are identified numerically from model 1 to model 50 on the vertical axis. Note that models are ordered such that first 25 models have “GMM-style” instruments while the second 25 models have “collapsed” instruments. Within the two groups, a model with the first to j^{th} lagged difference of log GDP per capita and the first to k^{th} lagged difference of the Gini coefficient as instruments is assigned model number $(j-1)5+k$, for $j,k=1,\dots,5$. So, for example, the model for the levels equation in which the instrument set consists of the first $j=2$ lagged changes of log income per capita and the first $k=3$ lagged changes in the Gini is assigned model number 8 when the instruments enter “GMM-style” with one instrument for each time period, and model number 25+8=33 when the instruments enter “collapsed”.

Several features of this graph are noteworthy. A first general observation is that the weak instrument robust confidence sets in the second, third, and fourth panels are very different from the conventional confidence intervals based on a standard Wald test shown in the first panel. Consistent with Table 6, the Wald confidence sets are all fairly small, and almost all of them include almost exclusively negative values for the estimated growth effect of inequality. In contrast, the weak instrument-consistent confidence sets are frequently quite large, particularly among the first 25 models which feature a large number of “GMM-style” instruments, as in the benchmark specifications of all of the papers reviewed here. While a few specifications deliver small confidence sets excluding zero, many of the confidence sets contain a wide range of positive and negative values of β. In fact, reading

vertically upwards from each hypothesized value of β on the horizontal axis, it is clear that every value of β between -2 and 2 is included in the weak instrument robust confidence set for at least one of the first 25 models.

The second key feature of this figure is that there is a stark difference between models that allow for "GMM-style" instruments versus "collapsed" instruments, i.e. the first 25 models versus the second 25 models. For models with "collapsed" instruments, the confidence sets are almost always empty when they are based on the AR and KJ tests. Recall that these test the joint null hypothesis that (a) the parameter is equal to its hypothesized value, and (b) the overidentifying restrictions are true at the hypothesized parameter values. In order to understand better which of these two hypotheses is being rejected throughout the parameter space, it is useful to unbundle the KJ test into its separate tests of these two null hypotheses. This reveals that in most cases the test rejects because the data strongly reject the overidentifying restrictions over the entire parameter space considered here. Finally, the CLR statistic tests only the null that the parameters are equal to their hypothesized values. Since this test assumes that the overidentifying restrictions are true, the confidence sets are not predominantly empty, as is the case for the AR and KJ-based confidence sets. However, it is interesting that the CLR-based tests include mostly positive values for the growth effect of inequality.

Overall, the picture that emerges from Table 6 and Figure 4 is discouraging for the robustness of conclusions that can be drawn from inequality and growth regressions estimated using the SYS-GMM estimator. On the one hand, a researcher relying only on conventional asymptotic approximations might point to the results in Table 5 and the first panel of Figure 4 to support the claim that SYS-GMM delivers a significantly negative estimated growth effect of inequality, that is robust to a wide range of possible choices of instruments. On the other hand, a researcher who takes instrument strength seriously and bases inferences about the estimated growth effect of inequality on the levels equation on the KJ test could find a combination of instruments that delivers a confidence set that (i) suggests a significant negative effect of inequality on growth, e.g. Model 3; (ii) suggests an "almost" significant positive effect on growth, e.g. Model 13; (iii) suggests an effect on growth that is insignificantly different from zero, e.g. most other models with "GMM-style" instruments, or (iv) that is empty, suggesting that the data reject the overidentifying restrictions at every hypothesized parameter value, e.g. most models with a "collapsed" instrument set.

Taken together, however, the overall picture that emerges from Figure 4 is fairly clear. Considering the weak instrument-consistent confidence sets in the second, third, and fourth panels, it is

clear that very few of them support strong conclusions about either a significantly positive or significantly negative effect of inequality on growth. However, the fact that there are a handful of cases that depart from this general pattern suggests that it is important for users of SYS-GMM to systematically document that their conclusions hold not just when appropriately accounting for instrument strength, but also across a range of possible instrument sets.

5. Conclusions

The question of the effect of inequality on growth is of central importance to researchers and policy makers. Several recent papers have contributed to an emerging empirical consensus that inequality is harmful for growth by reporting SYS-GMM estimates of cross-country growth regressions that show a statistically-significant negative effect of inequality on growth. However, an underappreciated concern with this empirical approach is that the internal instruments relied on by this estimator to isolate causal effects can be weak. A large literature on weak instruments has shown that when instruments are weak, inferences based on conventional Wald statistics are misleading. This literature has also developed alternative approaches to inference that are valid when instruments are weak. However, these techniques are rarely applied in cross-country growth empirics that rely on the SYS-GMM estimator, including the four studies reviewed here.

In this paper, I have used standard diagnostic tools from the literature on weak instruments to show pervasive evidence of weak internal instruments in the benchmark specifications of three of these inequality-and-growth studies. I also showed that weak instrument-consistent confidence sets for the estimated effect of inequality on growth include a wide range of positive and negative values, suggesting that strong conclusions about either negative or positive growth effects of inequality cannot be drawn from these benchmark specifications.

Of course, one should not generalize from this particular finding in a few benchmark empirical specifications in a small number of recent papers to the broader conclusion that cross-country regressions are necessarily uninformative about the effects of inequality on growth. One should also not conclude from this that the popular SYS-GMM estimator is not a useful tool for conducting cross-country growth empirics. Rather, this paper simply demonstrates that, once instrument strength is properly taken into account, the benchmark specifications in these recent empirical studies are not very informative about the effects of inequality on growth. This clearly does not rule out the possibility that there may be other empirical specifications for the relationship between inequality and growth that

support stronger conclusions -- in either direction -- on this important issue. Similarly, there may very well be other applications of SYS-GMM in cross-country growth empirics in which weak instrument problems do not arise, or even if they do, that weak instrument-consistent confidence sets nevertheless are meaningfully small and support strong conclusions about the sign and size of the growth effect of the variable under consideration.

More generally, this paper underscores the importance of routinely documenting instrument strength in applications of the SYS-GMM estimator to cross-country growth empirics, and of relying on weak instrument-robust approaches to inference as appropriate. Moreover, the results in the last section of this paper suggest the importance of also systematically documenting the sensitivity of conclusions about parameters of interest to a reasonably broad set of choices for constructing the instrument set within SYS-GMM. While users of these techniques should probably avoid using this sensitivity analysis to "mine" the universe of possible instrument sets to isolate those that are strong, the results of Section 4 do suggest that taking care to limit the proliferation in the number of instruments (by limiting lags and by "collapsing" the instruments) is likely to lead to stronger identification of the parameters of interest.[14] There probably also is scope to apply recent techniques from the literature on many weak instruments to the specific setting of the SYS-GMM estimator to improve instrument strength as well as to remove some of the discretion in the choice of the instrument set (see for example Hansen and Kozbur (2014) and references therein). Finally, consumers of cross-country empirical studies using the SYS-GMM estimator should expect authors to report weak instrument diagnostics, weak instrument-consistent approaches to inference, and sensitivity analysis over alternative instrument sets, in order to help them to assess the reliability of the evidence reported in these studies.

References

Anderson, T. W. and H. Rubin (1949). "Estimation of the Parameters of a Single Equation in a Complete Set of Stochastic Equations". The Annals of Mathematical Statistics". 20:46-63.

Arellano, Manuel and Stephen Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations". Review of Economic Studies. 58(2):277-97.

[14] This recommendation echoes a similar point made in Roodman (2009), but from the perspective of avoiding the risk of overfitting in the first stage regressions when the number of instruments is large relative to the number of observations.

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

Barro, Robert (2000). "Inequality and Growth in a Panel of Countries". Journal of Economic Growth. 5(1):5-32.

Baum, Christopher, Mark Schaffer, and Stephen Stillman (2007). "Enhanced Routines for Instrumental Variables/GMM Estimation and Testing". Stata Journal. 7(4):465-506.

Bazzi, Samuel and Michael Clemens (2013). "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth". American Economic Journal: Macroeconomics. 5(2):152-186.

Blundell, Richard and Stephen Bond (2000). "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models". Journal of Econometrics. 87(1):115-143.

Bun, Maurice and Frank Windmeijer (2009). "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models". The Econometrics Journal. 13:95-126.

Caselli, Francesco, Gerardo Esquivel and Fernando Lefort (1996). "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics". Journal of Economic Growth. 1(3):363-89.

Castelló-Climent, Amparo (2010). "Inequality and Growth in Advanced Economies: An Empirical Investigation". Journal of Economic Inequality. 2010(8):293-321.

Castelló-Climent, Amparo and Rafael Domenech (2002). "Human Capital Inequality and Economic Growth: Some New Evidence". Economic Journal. 112:C187-C200.

Chernozhukov, Viktor and Christian Hansen (2008). "The Reduced Form: A Simple Approach to Inference with Weak Instruments". Economics Letters. 100(1):67-71.

Dabla-Norris, Era, Kalpana Kochhar, Frantisek Ricka, Nujin Suphaphiphat, and Evridiki Tsounta (2015). "Causes and Consequences of Income Inequality: A Global Perspective". IMF Staff Discussion Note No. 15/13.

Dollar, David and Aart Kraay (2003). "Institutions, Trade and Growth: Revisiting the Evidence". World Bank Policy Research Department Working Paper No. 3004.

Easterly, William (2007). "Inequality Does Cause Underdevelopment: Insights from a New Instrument". Journal of Development Economics. 84(2):755-766.

Finlay, K., Magnusson, L.M., Schaffer, M.E. (2013). "weakiv: Weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models". http://ideas.repec.org/c/boc/bocode/s457684.html.

Forbes, Kristin (2000). "A Reassessment of the Relationship Between Inequality and Growth". American Economic Review. 90(4):869-87.

Halter, Daniel, Manuel Oechslin, and Josef Zweimüller (2014). "Inequality and Growth: The Neglected Time Dimension". Journal of Economic Growth. 19:81-104.

Hansen, Christian and Damian Kozbur (2014). “Instrumental Variables Estimation with Many Weak Instruments Using Regularized JIVE”. Journal of Econometrics. 182:290-308.

Holtz-Eakin, Douglas, Whitney Newey and Harvey S. Rosen (1988). “Estimating Vector Autoregressions With Panel Data”. Econometrica. 56(6): 1371-1395.

Jenkins, Stephen (2014). “World Income Inequality Databases: An Assessment of the WIID and SWIID”. Institute for Social and Economic Research Working Paper No. 2008-31.

Kleibergen, Frank (2002). “Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression”. Econometrica. 70(5):1781-1803.

Kleibergen, Frank (2004). “Testing Subsets of Structural Parameters in the Instrumental Variables Regression Model”. Review of Economics and Statistics. 86:418-423.

Kleibergen, Frank (2005). “Testing Parameters in GMM Without Assuming that They Are Identified”. Econometrica. 73:1103-1123.

Kleibergen, Frank and Richard Paap (2006). “Generalized Reduced Rank Tests Using the Singular Value Decomposition”. Journal of Econometrics. 133:97-126.

Mikusheva, Anna (2013). “Survey on Statistical Inferences in Weakly-Identified Instrumental Variables Models”. Applied Econometrics. 29(1):117-131.

Moreira, Marcelo (2003). “A Conditional Likelihood Ratio Test for Structural Models”. Econometrica. 71:1027-1048.

Murray, Michael (2006). "Avoiding Invalid Instruments and Coping with Weak Instruments". Journal of Economic Perspectives. 20(4):111-132.

Nelson, Charles and Richard Startz (1990). “Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator”. Econometrica. 58(4):967-976.

Ostry, Jonathan, Andrew Berg, and Charalambos Tsangarides (2014). “Redistribution, Inequality, and Growth”. IMF Staff Discussion Note SDN14/02.

Roodman, D. (2009). “How to do xtabond2: An Introduction to Difference and System GMM in Stata”. Stata Journal 9(1): 86-136.

Sanderson, Eleanor and Frank Windmeijer (2015). “A Weak-Instrument F-test in Linear IV Models with Multiple Endogenous Variables”. Journal of Econometrics. *In press.*

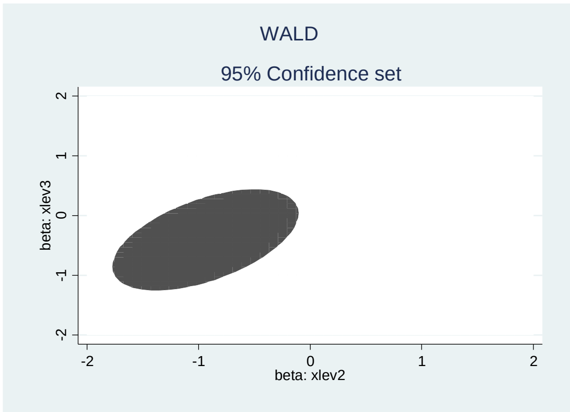
Solt, F. (2009). “Standardizing the World Income Inequality Database”. Social Science Quarterly. 90(2):231–42.

Stock, James and Motohiro Yogo (2005). “Testing for Weak Instruments in Linear IV Regression”, in *Identification and Inference for Econometric Models: Essays in Honour of Thomas Rothenberg*, edited by Donald Andres and James Stock. 80-108. New York: Cambridge University Press.

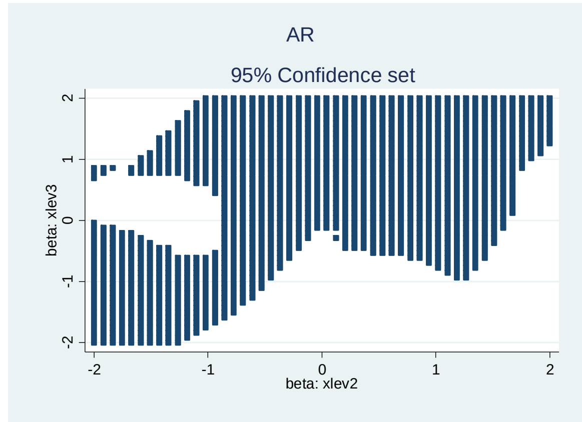
Voitchovsky, Sarah (2005). "Does the Profile of Income Inequality Matter for Economic Growth? Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution". Journal of Economic Growth. 10(3):273-296.

Figure 1: Confidence Sets for the Coefficient on Inequality (Horizontal Axis) and Redistribution (Vertical Axis) in the Levels Equation -- Ostry, Berg and Tsangarides (2014) Benchmark Specification

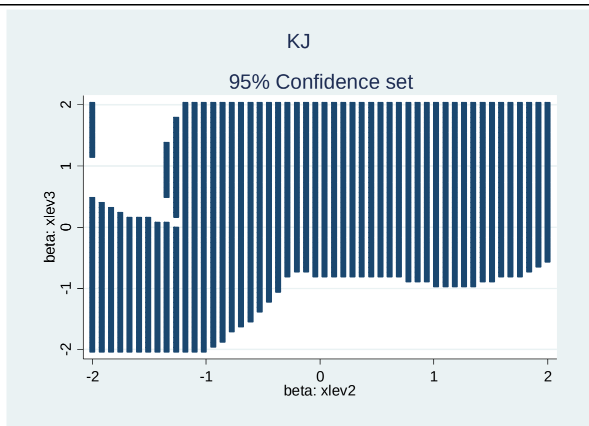
Panel 1: Based on Wald Statistic



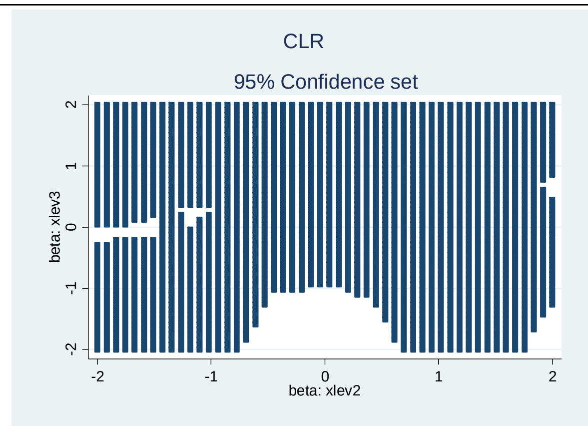
Panel 2: Based on AR-Statistic



Panel 3: Based on KJ Statistic



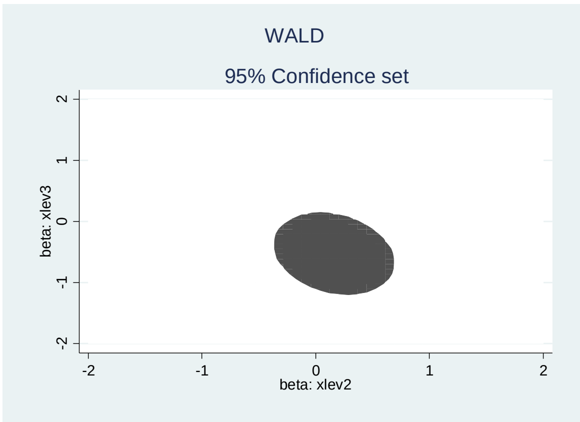
Panel 4: Based on CLR Statistic



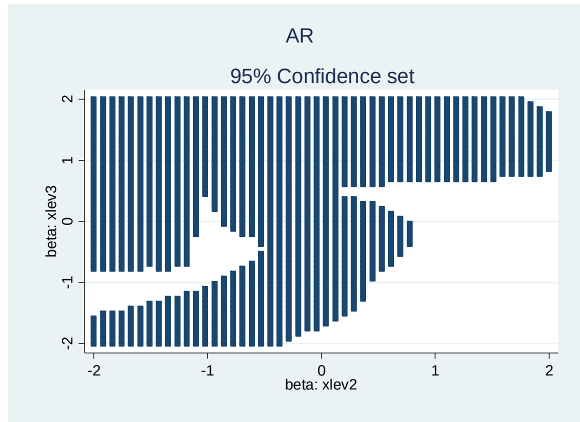
Notes: This figure reports confidence sets for the coefficient on the net gini (horizontal axis) and redistribution (vertical axis). The dark-shaded region corresponds to combinations of hypothesized values for these two parameters that jointly are not rejected by the Wald, AR, KJ and CLR tests.

Figure 2: Confidence Sets for the Coefficient on Lagged Inequality (Horizontal Axis) and Twice-Lagged Inequality (Vertical Axis) in the Levels Equation -- Halter, Oechshlin and Zweimüller (2014) Benchmark Specification

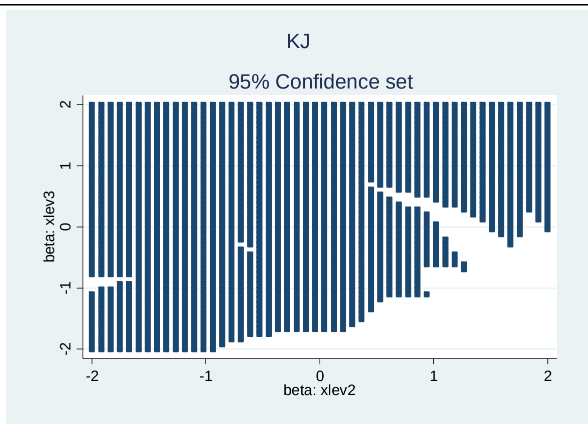
Panel 1: Based on Wald Statistic



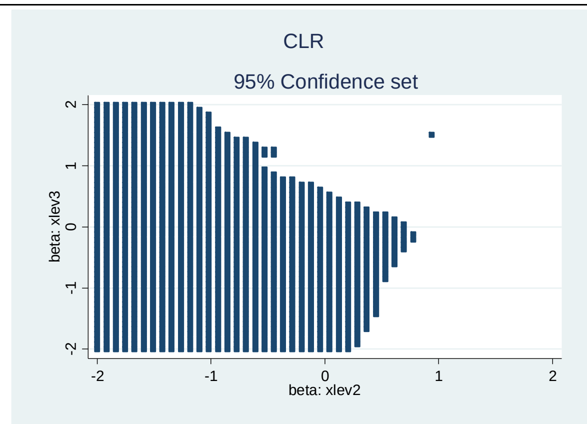
Panel 2: Based on AR-Statistic



Panel 3: Based on KJ Statistic



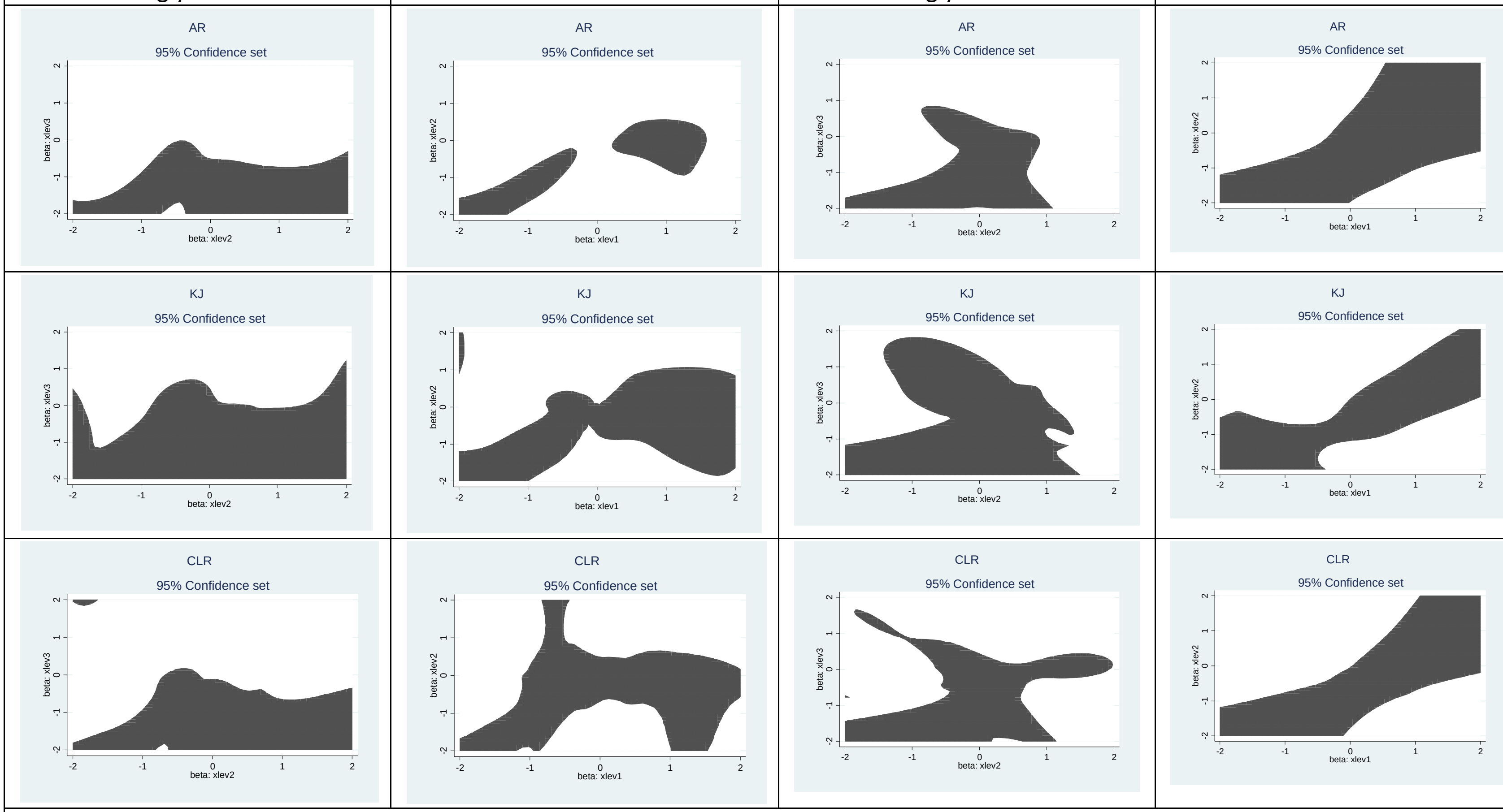
Panel 4: Based on CLR Statistic



Notes: This figure reports confidence sets for the coefficient on the lagged gini (horizontal axis) and twice lagged gini (vertical axis). The dark-shaded region corresponds to combinations of hypothesized values for these two parameters that jointly are not rejected by the Wald, AR, KJ and CLR tests.

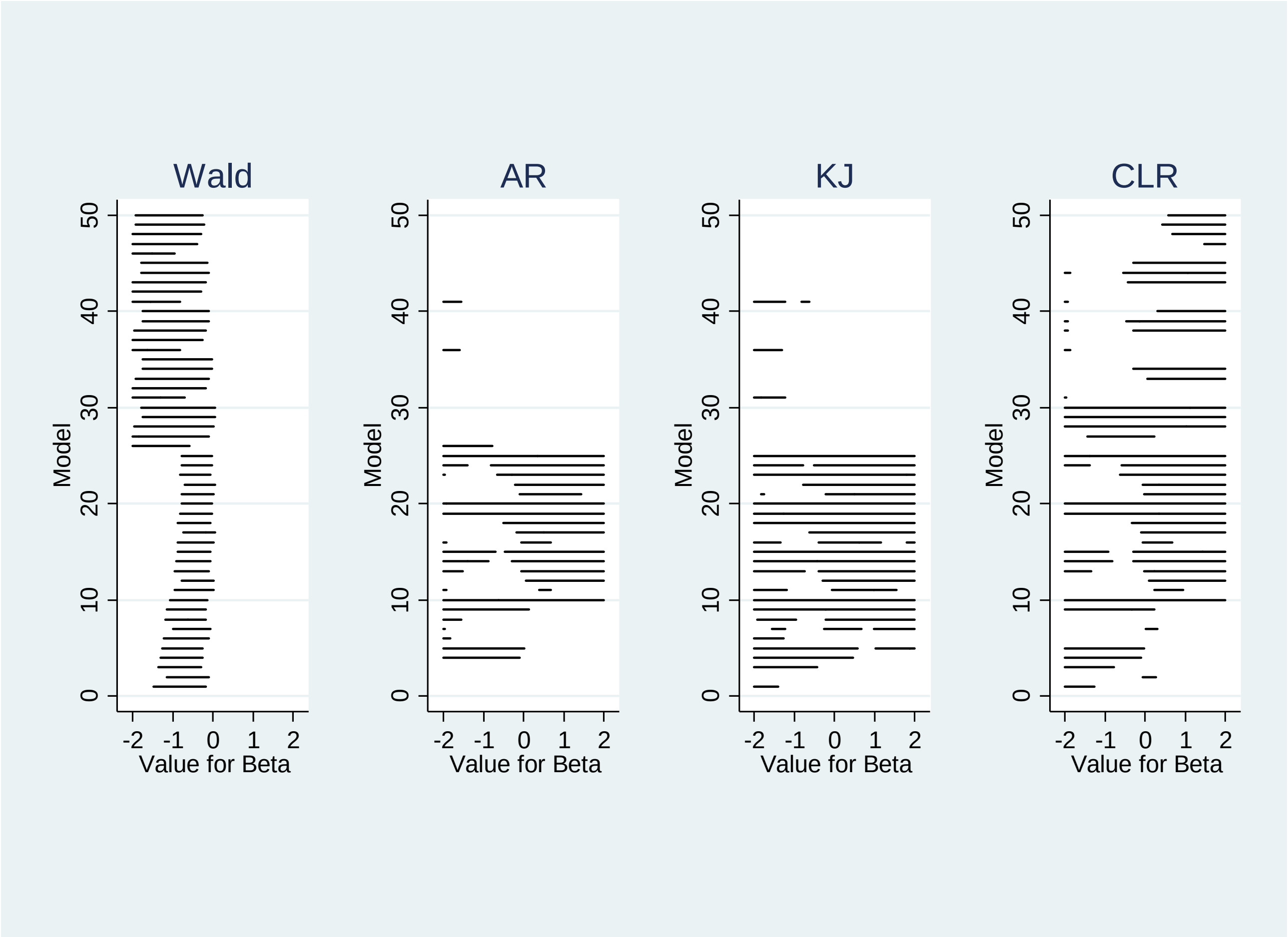
Figure 3: Alternative Two-Dimensional Confidence Sets

OBT -- Coefficient on Initial Income is:		HOZ-- Coefficient on Initial Income is:	
Strongly Identified	Fixed at SYS-GMM Estimate	Strongly Identified	Fixed at SYS-GMM Estimate



Notes: This figure reports confidence sets for the coefficients on the two inequality variables of interest in OBT and HOZ. For OBT, the graphs show the confidence sets for the net gini (horizontal axis) and redistribution (vertical axis). For HOZ, the graphs show the confidence sets for lagged inequality (horizontal axis) and twice-lagged inequality (vertical axis). The dark-shaded region corresponds to combinations of hypothesized values for these two parameters that jointly are not rejected by the Wald, AR, KJ and CLR tests. In the first and third columns the coefficient on initial income is assumed to be strongly identified, while in the second and fourth columns it is fixed at its point estimate from the SYS-GMM estimator.

Figure 4: Conventional and Weak Instrument-Robust Confidence Sets for Growth Effect of Inequality Based on the Levels Equation



Notes: This figure reports confidence sets for the estimated growth effect of inequality from the levels regression of a variant of OBT that excludes redistribution from the growth regression. Each horizontal line represents a different model with a different choice of lags of changes in income and inequality used as internal instruments. In the first 25 models the instruments enter "GMM style" with one variable per time period for each instrument. In the second 25 models the instruments are "collapsed" with one column in the instrument matrix for each instrument.

Table 1: Summary of Inequality and Growth Studies

Paper	*Choice of Explanatory Variables*	*Location of Benchmark Specification*	*Source of Inequality Data*
Ostry, Berg, and Tsangarides (2014) (OBT)	$\beta' X_{it-1} = \beta_1 NetGini_{it-1} + \beta_2 Redistribution_{it-1}$	Table 3, Column 1	Solt (2009)
Halter, Oechslin, and Zweimüller (2014) (HOZ)	$\beta' X_{it-1} = \beta_1 Gini_{it-1} + \beta_2 Gini_{it-2} + \beta_3 Schooling_{it-1}$ $+\beta_4 Investment_{it-1} + \beta_5 Price\ Level\ of\ Investment_{it-1}$	Table 3, Column 4	UNU-Wider World Income Inequality Database
Dabla-Norris, Kochhar, Suphaphihpat, Ricka and Tsounta (2015) (DKSRT)	$\beta' X_{it-1} = \beta_1 NetGini_{it-1} + \beta_2 \Delta y_{it-1}$	Table 1, Column 1	Solt (2009)
Castelló-Climent (2010) (CC)	$\beta' X_{it-1} = \beta_1 Gini_{it-1} + \beta_2 Schooling_{it-1}$ $+\beta_3 Government\ Spending/GDP_{it-1} + \beta_4 Trade_{it-1}$ $+\beta_5 Inflation_{it-1}$	Table 1, Column 1	Human capital inequality from Castelló and Domenech (2002)

Notes: This table summarizes the choice of explanatory variables included in the benchmark specifications of the four inequality and growth studies considered here. These choices are embedded in the dynamic panel growth regression in Equation (1) in the text.

Table 2: Ostry, Berg and Tsangarides (2014) Benchmark Specification

	(1)	(2)	(3)
	SYS-GMM	Differenced Equation, 2SLS	Levels Equation, 2SLS
Lagged Income	0.970*** (0.017)	0.281*** (0.108)	0.973*** (0.019)
Lagged Net Gini	-0.708*** (0.215)	0.525 (0.399)	-0.939*** (0.297)
Lagged Redistribution	-0.033 (0.259)	0.540 (0.485)	-0.409 (0.300)
P-value for Ho: Overidentification	0.300	0.117	0.036
P-value for Ho: Underidentification		0.177	0.168
Cragg-Donald Wald F-Statistic		2.03	2.01
Crit. Val. for 30% Maximal Relative Bias		3.93	4.13
Sanderson-Windmeijer Conditional F-statistics for lagged Gini (Redistribution)		20.62 (32.59)	8.25 (7.49)
Crit. Val. for 25% Maximal Size		49.22	26.08
Weak Instrument-consistent 95% Confidence Sets Based on Projecting AR-Based Confidence Sets			
Net Gini		[… , …]	[… , …]
Redistribution		[… , …]	[… , …]
Weak-Instrument-Consistent 95% Confidence Sets Based on Projecting KJ-Based Confidence Sets			
Lagged Net Gini		[… , …]	[… , …]
Lagged Redistribution		[… , …].	[… , …]
Weak Instrument-Consistent 95% Confidence Sets Based on Projecting CLR-Based Confidence Sets			
Lagged Net Gini		[… , …]	[… , …]
Lagged Redistribution		[… , …]	[… , …]
Period Dummies	Yes	Yes	Yes
Instruments In Differenced Equation	$y_{it-2}, \dots, y_{it-5}$ G_{it-2}, G_{it-3}, R_{it-2}, R_{it-3}	$y_{it-2}, \dots, y_{it-5}$ G_{it-2}, G_{it-3}, R_{it-2}, R_{it-3}	
Instruments In Levels Equation	$\Delta G_{it-2}, \Delta G_{it-3}$, $\Delta R_{it-2}, \Delta R_{it-3}$		$\Delta G_{it-2}, \Delta G_{it-3}$, $\Delta R_{it-2}, \Delta R_{it-3}$
Number of Excluded Instruments	106	72	34
Number of Countries	130	125	130
Number of Observations	828	678	828

Table 3: Halter, Oechslin and Zweimüller (2014) Benchmark Specification			
	(1)	(2)	(3)
	SYS-GMM	Differenced Equation, 2SLS	Levels Equation, 2SLS
Lagged Income	0.998*** (0.013)	-0.040 (0.278)	0.999*** (0.015)
Lagged Gini	0.160 (0.178)	1.897*** (0.661)	0.079 (0.201)
Twice-Lagged Gini	-0.671*** (0.206)	0.939*** (0.201)	-0.754*** (0.239)
P-value for Ho: Overidentification	0.022	0.023	0.158
P-value for Ho: Underidentification		0.140	0.205
Cragg-Donald Wald F-Statistic		1.50	1.67
Crit. Val. for 30% Maximal Relative Bias		4.59	4.39
Sanderson-Windmeijer Conditional F-statistics for lagged Gini (twice lagged Gini)		4.15	23.46 (10.14)
Crit. Val. for 25% Maximal Size		12.82	14.00
Weak Instrument-consistent 95% Confidence Sets Based on Projecting AR-Based Confidence Sets			
Lagged Gini		[0.94, ...]	[... , ...]
Twice-Lagged Gini		[-0.20, 1.35]	[... , ...]
Weak-Instrument-Consistent 95% Confidence Sets Based on Projecting KJ-Based Confidence Sets			
Lagged Gini		[0.37, ...]	[... , ...]
Twice-Lagged Gini		[-1.18, 1.84]	[... , ...]
Weak Instrument-Consistent 95% Confidence Sets Based on Projecting CLR-Based Confidence Sets			
Lagged Gini		[0.78, ...]	[... , 0.78] U 0.94
Twice-Lagged Gini		[-0.04, 1.27]	[... , ...]
Period Dummies	Yes	Yes	Yes
Excluded Instruments In Differenced Equation	y_{it-2}, G_{it-2}	y_{it-2}, G_{it-2}	
Excluded Instruments In Levels Equation	$\Delta y_{it-1}, \Delta G_{it-1}$		$\Delta y_{it-1}, \Delta G_{it-1}$
Number of Excluded Instruments	26	12	14
Number of Countries	88	62	88
Number of Observations	286	191	286
Note: Included instruments are period effects in both the levels and differenced equations, and ΔG_{it-2} in the differenced equation.			

Table 4: Castelló-Climent (2010) Benchmark Specification

	(1)	(2)	(3)
	SYS-GMM	Differenced Equation, 2SLS	Levels Equation, 2SLS
Lagged Income	0.978*** (0.035)	0.345* (0.161)	0.941*** (0.073)
Lagged Gini	-0.245 (0.160)	0.817 (0.510)	-0.505 (0.381)
P-value for Ho: Overidentification	0.321	0.477	0.06
P-value for Ho: Underidentification		0.022	0.044
Cragg-Donald Wald F-Statistic		2.96	2.06
Crit. Val. for 30% Maximal Relative Bias		4.11	4.48
Sanderson-Windmeijer Conditional F-statistic for lagged Gini		13.29	2.04
Crit. Val. for 25% Maximal Size		30.94	15.19
Weak Instrument-consistent 95% Confidence Sets Based on AR Statistic			
Lagged Gini		[-1.27, 1.68]	[-1.27, -0.10]
Weak-Instrument-Consistent 95% Confidence Sets Based on KJ Statistic			
Lagged Gini		[… ,-1.27] U [-1.11 , 1.07]	[-1.52, 0.06]
Weak Instrument-Consistent 95% Confidence Sets Based CLR Statistic			
Lagged Gini		[-0.75, 1.07]	[-1.19,-0.06]
Period Dummies	Yes	Yes	Yes
Excluded Instruments In Differenced Equation	$y_{it-2}, \dots, y_{it-4}, G_{it-2}, \dots, G_{it-4}$	$y_{it-2}, \dots, y_{it-4}, G_{it-2}, \dots, G_{it-4}$	
Excluded Instruments In Levels Equation	$\Delta y_{it-1}, \Delta G_{it-1}$		$\Delta y_{it-1}, \Delta G_{it-1}$
Number of Excluded Instruments	58	42	16
Number of Countries	102	101	102
Number of Observations	809	707	809

Note: Period effects are included instruments.

Table 5: Weak Instrument-Consistent Confidence Sets for Levels Equation -- Alternative Treatments of Coefficient on Initial Income

	Coefficient on Initial Income is:	
	Strongly Identified	**Fixed at SYS-GMM Estimate**
OBT: Lagged Inequality		
AR	[… , …]	[… ,-0 .37] U [0.29,1.51]
KJ	[… , …]	[… , …]
CLR	[… , …]	[… , …]
OBT: Lagged Redistribution		
AR	[… , -0.04]	[… , 0.53]
KJ	[… , 1.18]	[… , …]
CLR	[… ,0.12] U [1.92,…]	[… , …]
HOZ: Lagged Inequality		
AR	[… , 1.02]	[… , …]
KJ	[… , 1.43]	[… , …]
CLR	[… , 1.92]	[… , …]
HOZ: Twice Lagged Inequality		
AR	[… , 0.78]	[… , …]
KJ	[… , 1.76]	[… , …]
CLR	[… , 1.67]	[… , …]
CC: Lagged Inequality		
AR	[-0.85, -0.31] U [0.02, …]	[-0.34, 0.22]
KJ	[-1.92,-1.59] U [-0.95, …]	[-0.39, 0.33]
CLR	[-0.76,-0.38] U [-0.31, …]	[-0.32, 0.22]

Note: This table reports weak instrument-consistent confidence sets for the coefficients on the main inequality variables of interest in OBT, HOZ and CC. In the case of OBT and CC, these come from projecting the confidence sets in Figure 3 onto the corresponding axes. For CC the confidence sets are exact. In the first column the coefficient on initial income is assumed to be strongly identified, while in the second column it is fixed at its point estimate from the SYS-GMM estimator.

Table 6: Robustness Across Alternative Instrument Sets
(OBT Dataset, Levels Equation, Redistribution Not Included)

Panel A -- SYS-GMM Estimates								
	N	**mean**	**sd**	**p10**	**p25**	**p50**	**p75**	**p90**
Coefficient on Lagged Inequality								
GMM-style	625	-0.55	0.13	-0.76	-0.66	-0.51	-0.44	-0.40
Collapsed	625	-1.18	0.33	-1.81	-1.20	-1.06	-0.97	-0.89
Both	1250	-0.86	0.40	-1.41	-1.06	-0.81	-0.51	-0.43
Absolute value of t-Statistic for Lagged Inequality								
GMM-style	625	2.49	0.35	2.05	2.23	2.45	2.74	2.99
Collapsed	625	2.85	0.48	2.30	2.52	2.77	3.07	3.63
Both	1250	2.67	0.46	2.13	2.35	2.62	2.92	3.25
Panel B -- Differenced Equation								
	N	**mean**	**sd**	**p10**	**p25**	**p50**	**p75**	**p90**
Coefficient on Lagged Inequality								
GMM-style	25	0.48	0.21	0.28	0.31	0.43	0.59	0.83
Collapsed	25	2.88	2.48	1.41	1.90	2.12	3.15	3.78
Both	50	1.68	2.13	0.30	0.43	0.96	2.12	3.38
Absolute value of t-Statistic for Lagged Inequality								
GMM-style	25	1.07	0.40	0.70	0.77	0.99	1.28	1.78
Collapsed	25	1.88	0.54	1.12	1.71	2.08	2.26	2.41
Both	50	1.48	0.62	0.70	0.91	1.40	2.08	2.32
Panel C -- Levels Equation								
	N	**mean**	**sd**	**p10**	**p25**	**p50**	**p75**	**p90**
Coefficient on Lagged Inequality								
GMM-style	25	-0.53	0.15	-0.80	-0.66	-0.47	-0.41	-0.37
Collapsed	25	-1.29	0.52	-2.29	-1.21	-1.07	-0.96	-0.89
Both	50	-0.91	0.54	-1.74	-1.07	-0.84	-0.47	-0.40
Absolute value of t-Statistic for Lagged Inequality								
GMM-style	25	2.77	0.65	1.96	2.49	2.71	2.96	3.94
Collapsed	25	2.29	0.37	1.79	1.99	2.24	2.53	2.87
Both	50	2.53	0.58	1.90	1.99	2.50	2.83	3.44

Note: This table summarizes the distribution of point estimates of the growth effect of inequality from 2SLS estimates of the levels equation, and the corresponding (absolute) t-statistics. These distributions are calculated over different instrument sets consisting of different sets of lags of changes in income and inequality used as internal instruments, as described in the text.