

# The Institutional Foundations of Inequality and Growth

Lewis Davis  
*Union College*  
[davisl@union.edu](mailto:davisl@union.edu)

and

Mark Hopkins  
*Gettysburg College*  
[mhopkins@gettysburg.edu](mailto:mhopkins@gettysburg.edu)

**Abstract:** One unresolved question in the empirical growth literature is the impact of income inequality on future economic growth. While a first wave of cross-country regressions concluded that there was a significant and *negative* impact on future growth, a second wave of regressions using panel data to control for potential omitted variable bias has consistently found the coefficient on inequality to be *positive*. A third wave of the literature has made methodological contributions, but has done little to offer a clear story regarding the economic relationships between these variables that might explain the evidence. This paper offers one such story, starting from two fairly uncontested claims: 1) growth and inequality statistics each summarize some aspect of the same underlying distribution dynamics, and thus will be determined jointly by those dynamics; and 2) institutions, particularly those which restrict economic participation and those that create investment incentives for that participation, are likely to be an important determinant of those distributional dynamics. We show that a simple empirical model with these features helps to reconcile the apparent “paradox” suggested by the first and second wave regressions.

**Draft:** July 25, 2006

## **1 Introduction**

Since the pioneering work of Barro (1991), empirical work on the determinants of economic growth have found strong evidence of conditional convergence, or the prediction from the neoclassical growth model that, *ceteris paribus*, a higher level of current per-capita income results in slower future growth. Much less consensus exists regarding what impact, if any, the *distribution* of current income has on growth rates. Both the informal and formal evidence appears to be contradictory.

One early motivation for research on this question was the observation that many (high inequality) Latin American and African countries have experienced disappointing growth, in part because distributional conflict has resulted in political uncertainty, populist policies and macroeconomic instability. Initial empirical research by Alessina and Rodrik (1994) and Persson and Tabellini (1994), concluded from cross-country growth regressions that inequality was indeed associated with slower future growth.

More recently, however, a competing story has been popularized in which inequality and growth are positively linked. Progressive social policies in general, and institutional inflexibility in the labor market in particular, has been cited as one reason why the EU has experienced both lower inequality and lower growth rates than in the United States. Furthermore, a second wave of empirical evidence has provided some formal support for the claim that inequality and growth may be positively linked. Partridge (1998), replicating Persson and Tabellini's methodology on U.S. states, finds a positive coefficient on initial inequality, as does Forbes (2000) and others using more recently available panel data across countries to control for country-specific fixed effects.

One might well ask, in the words of Bannerjee and Duflo (2003), “Is there anything then, apart from the obvious fact of disagreement, that we can take away from this body of evidence?” The answer we suggest in this paper is *yes*: the results of the literature to date may not be contradictory evidence on one relationship, but rather consistent evidence on two. In particular, at least two things are noteworthy about the literature to date. First, most economic models do not suggest a direct causal role of inequality on growth; rather, they suggest either an indirect role acting through market, political or social mechanisms, or no role at all: inequality and growth may each be simply features of the same underlying income distribution dynamics, jointly determined by some underlying variables. As a result, the policy conclusions that can be taken away from the current crop of reduced form growth-inequality regressions are limited.

Second, the main variables determining the relationship between inequality and growth in theory – the quality of the credit markets, the level of democracy, the role of government in markets – are all what might be characterized as “institutional” variables. Hall and Jones (1999), Acemoglu, Johnson and Robinson (2001, 2002) and others have provided strong evidence that institutional quality matters for the level of average income. If this is true, it is likely that they also matter for the distribution of income. We believe that acknowledging and identifying these institutional variables is essential both for gaining proper insight into the relationship between inequality and growth, as well as consistent and efficient estimation of that relationship.

Davis (2006) uses a two sector growth model with positive externalities in one (“formal”) sector to reproduce the main stylized facts on the growth-inequality nexus suggested by the literature to date. In that model growth and inequality are endogenously

determined according to two parameters related to institutional structure – barriers to participation in the formal sector, which are relatively time invariant and have opposite effects on growth and inequality; and a time-varying level of government redistribution, which drives growth and inequality in similar directions.

We use this paradigm to motivate our empirical approach in this paper, examining the degree to which relatively time-invariant institutional barriers to formal sector participation may be the key omitted variable driving the disparity between cross-sectional and panel data estimates of the growth-inequality relationship. We find that measures of expropriation risk and democracy are important in explaining both growth and inequality, and that after controlling for these in a cross-sectional growth regression the coefficient on inequality ceases to be negative and significant. Noting, however, that in the long-run growth and inequality are likely to be co-determined features of the underlying data-generating income distribution dynamics, and that growth and inequality are also both likely to drive the quality of institutional development as well, we employ a number of instrumenting techniques to control for potential endogeneity problems. Our results appear to be largely robust using a standard set of instruments.

Finally, we turn to the question of efficient estimation. As we confirm, cross-country regressions of growth on inequality are subject to potential omitted variable bias. Fixed-effects estimates from panel data are highly inefficient, however, because inequality varies relatively little within countries (relative to its variation across countries).<sup>1</sup> Furthermore, country-invariant characteristics can not be identified in such a regression as they are swept up in the country dummy, which means that omitted variable bias can be avoided, but the reasons for it not identified. The efficient GLS estimator, or

---

<sup>1</sup> Furthermore, Barro (2000) has argued that much of this variation may in fact be measurement error.

random-effects estimator, does allow conditioning on country-invariant characteristics (which include many measures of institutional quality) but it requires a strong exogeneity assumption the any included fixed variables are uncorrelated with those excluded. By including what we believe to be the key omitted variables and using the Hausman-Taylor (1981) estimator to instrument for any remaining endogeneity among the included variables, we attempt to measure the impact of inequality on growth more efficiently than previous work.

In summary, the goal of this paper is to examine the institutional determinants of inequality and growth to ascertain *whether* institutions can explain the apparent conflict in the evidence, *which* institutions in particular and, if so, *what* are the “real” effects of inequality on growth, with the ultimate goal of clarifying what policy conclusions – if any – can be taken away from this literature.

## **2 A Review of Literature on Growth and Inequality**

Many theories have been proposed in which inequality has a causal effect on growth, although the nature and direction of the effect varies a great deal. In much of the post-Keynesian literature differential saving rates are assumed to exist among the rich and poor, suggesting that average investment will rise along with the concentration of income among the rich. If expected future consumption is considered a luxury good, this assumption certainly seems plausible. When capital markets are imperfect, however, Loury (1978), Galor and Zeira (1993) and others have argued that concentration of income may lead to *lower* growth as more households find themselves credit-constrained

and unable to undertake profitable investments. On the other hand, Barro (2000) and Banerjee and Duflo (2003) have noted that if the investment technology entails a significant fixed cost, a mean-preserving spread in the distribution of income among poorer countries may lead to higher growth as a higher fraction of households find themselves above the threshold necessary to finance the investment. Persson and Tabelini (1994) and Alesina and Rodrik (1994) both argue that distribution can have political consequences and, among democracies in particular, greater inequality may increase pressure for distortionary redistribution.

Perotti (1996) and Benabou (1996) review much of this literature and test a number of these theories, concluding that overall the relationship between inequality and growth across countries appears to be negative. This result may reflect the influence of certain country characteristics omitted from the regression, however. One can imagine a number of economic characteristics associated with the degree to which economic opportunity is accessible to all, which would be associated both with higher growth and lower inequality. Due to the heterogeneity of country observations and the vast array of potential growth determinants, it is often unclear whether the conditioning in parsimonious cross-country regressions is sufficient to ensure correct parameter inference.

Two approaches can be taken to address this concern: one can either look for observations on a more homogeneous set of polities or adopt more sophisticated conditioning techniques. Partridge (1998) adopts the first approach, looking at the relationship of inequality and growth over U.S. states, which offers a much more homogeneous set of observations, while a number of other authors including Forbes(2000), Benhabib and Spiegel (1998), Li & Zou (1998) and Lundberg & Squire

(2003) adopt the second, exploiting panel data on inequality compiled by Deininger and Squire (1996) to control for potential omitted variable bias through country-specific fixed effects. In each case, the authors in what we will henceforth call this “second-wave” of the literature find evidence of a positive coefficient on inequality.

Robert Barro (2000), using a random-effects estimator, finds the impact of inequality to be relatively insignificant overall. The result is not surprising, perhaps, given that the random effects estimator is constructed as a weighted average of the between- and within- country estimates. Barro highlights one interesting feature of the results, however: breaking down countries by income level, the relationship appears to be slightly negative among poor countries and slightly positive among rich countries. While the explanation for this is as yet uncertain, Barro notes that this may help to explain the divergence between the negative cross-sectional and positive panel data estimates: poor countries are more numerous, and thus more highly represented in a cross-sectional sample while rich countries tend to have more complete data, and thus are more highly represented in the panel data samples – particularly those using dynamic panel data estimators requiring information on past lags.

Recently, in what might be considered a third-wave of the literature, the very use of such simple reduced form relationships has been called into question. Barro (2000) and Lundberg and Squire (2005?) both note that inequality itself is itself likely to be endogenous, and estimate growth and inequality jointly using a system of equations.<sup>2</sup> Banerjee and Duflo (2003) argue that all of these regressions are misspecified, however, and that the true underlying relationship is not between the *level* of inequality and growth,

---

<sup>2</sup> Barro also includes an equation for investment, to separate the direct impact of inequality on growth from the indirect impact through investment.

but *changes* in inequality and growth. They note that the hypothesized negative relationship between the level of inequality and future growth suggested by political economic models is through a causal chain linking (a) higher levels of inequality to (b) greater pressure for redistribution to (c) lower growth. They argue that the only clear structural relationship in the data is from (b) to (c), and that this relationship is, in fact, non-linear. They suggest that this may explain variation in the estimates found from direct regressions of the level of inequality on growth. Their conclusion is that caution must be used in drawing inference from the previous literature – making it an important addition to the literature, but one focused more on estimation and the mechanics of distributional changes than a clear story regarding patterns of economic development.

### ***Our empirical methodology***

Our empirical work employs a panel of data for 6 five-year periods starting in 1961-65 and ending in 1996-2000. Using this data, Table 1 replicates the basic stylized facts found in the literature described above using the specification employed by Forbes (2000), following Perotti (1996).<sup>3</sup> Growth in each period is regressed on the log of per-capita income and secondary years of schooling taken from the final year of the previous period, an income Gini coefficient around the final year of the previous period, and the

---

<sup>3</sup> In our regression we use a single composite measure of secondary years of schooling rather than allowing male and female secondary schooling to enter separately for two reasons. First, using their specification a hypothesis test that these coefficients were equal could not be rejected ( $p=64\%$ ), which motivates parsimony for its own sake. More importantly, however, we see no a priori reason to believe that the direct productivity impact of secondary schooling should differ significantly for men and women, and that differences in the reduced form estimates are most likely picking up some institutional characteristics which we hope to characterize explicitly in this paper. It should also be noted that the fixed effects constructed from our parsimonious specification, which we use employ in later analysis, have an extremely high correlation (over 0.97) with those which would be constructed from the original Forbes (2000) specification.

average price of investment (*PPP*) over the previous period. Data on per-capita income and the price of investment is taken from the Penn World Tables (version 6.1), data on years of schooling from Barro and Lee (2000), and data on income Gini coefficients was taken from the World Income Inequality Database (WIID), version 2 (2003), which contains as a subset the Deininger-Squire database used by Forbes(2000) and others. Because of marked variations in survey type and sample coverage, a number of adjustments were made to the WIID data to attempt to improve consistency across observations by controlling for differences in survey type and year of measurement.<sup>4</sup>

The first column of Table 1 replicates the conclusion of Forbes (2000) and others in the second-wave using fixed effects estimators that, within countries, increases in inequality have a positive and statistically significant impact on inequality. The economic impact is also significant, although about 2/3 of the magnitude of the effect claimed by Forbes (2000): we find a one standard deviation change in inequality within countries (roughly 3.4 points on the 100 point Gini scale) is associated with an increase in the average growth rate of 0.27 percentage points, or about 14% of a standard deviation of within-country growth rates. Column two uses a between estimator (regressing country average growth rates on country averages of inequality and the other variables), to replicate the findings of the original first wave results that *across* countries an increase

---

<sup>4</sup> Gini coefficients constructed based on income surveys are typically much higher than Gini coefficients constructed from consumption and expenditure surveys. Variation also exists in area coverage (whether rural or urban only), population coverage, and the survey unit (whether a household or an individual). Hedonic regressions were used to estimate the impact of these factors, exploiting variation across survey types within the same country and year. The Gini coefficients were then adjusted to represent gross income sampled over all areas by household and adjusted to a per-capita basis, in line with recommendations of the Canberra group. When data was not available for the specific year in question, Gini coefficients were used according to the following rule: take the closest available year, favoring past to future observations, up to three years back. If this resulted in a missing observation, but multiple data was available within 10 years a value for the target year was estimated as a linear interpolation of the available data. The net result was a panel of 623 Gini coefficients. Forbes, by contrast, uses 180 coefficients drawn from the Deininger and Squire (1996) database. The Gini dataset is available from the authors.

in inequality is associated with a *reduction* in growth. In particular, a one standard deviation in inequality (roughly 10 points on the Gini scale) decreases average growth by approximately 0.41 percentage points, or 21% of the standard deviation of growth across countries. In other words, although the direct effect of inequality appears to be larger in magnitude within countries than across them, the fact that inequality has fairly little time series variation actually means the role of inequality is somewhat more economically significant in explaining growth across countries than within them.

The third column presents results the efficient GLS estimator, using random effects. The GLS estimator is computed using a weighted average of the between and within estimates, with the weight derived from the relative importance of the country-specific error component in the overall variance. Because omitted country-specific effects appear to be relatively important, the GLS estimator more closely reflects the between-country estimates – although in this case the coefficient on inequality is no longer significantly different from zero. This finding is consistent with those of Barro (2000) and Forbes (2000) using random effects.

So how should we interpret these basic stylized facts? There are a number of mechanical explanations that may explain the dichotomy of results between short- and long-run effects, and between- and within- countries [*some of which are discussed thoroughly by Banerjee and Duflo (2003) and not repeated here, and others which I've cut out and now appear in an Appendix A*]. Our goal in this paper is not enumerate all of these possibilities, however, but to set forth a simple story of economic and institutional development that is consistent with the basic stylized facts established by the empirical literature. The principal components of this story are as follows:

- 1) Growth and inequality are each summary statistics derived from the underlying dynamics of the income distribution. As a result, they are likely to be jointly determined by many of the same underlying economic fundamentals. We believe that many of these economic fundamentals are, in turn, products of the country's institutional environment.
- 2) Institutional characteristics vary both across countries and across time, but those principally important for determining the long-run levels of income and inequality are likely to be relatively invariant over time. In particular, institutions that act to restrict opportunity across a broad range of the population – whether through *de facto* lack of access to education, land, credit or to the political process itself, or through policies that impose barriers to broad participation in the formal sector created by regulatory costs of entry for capital and labor – are likely to raise inequality while lowering growth.
- 3) Omitting such variables in a cross-country regression of growth on inequality will result in a negative bias on the estimated coefficient on inequality, and could explain the different results between the first and second wave of the literature. Although using country fixed-effects will eliminate this bias, it will result in inefficient estimates (relative to random effects) and potentially misleading inference (in confusing demand- with supply-driven growth in the relatively short, high frequency growth periods).
- 4) Institutions do vary within countries, although in this case that variation generally involves public policy more than deep changes in the underlying institutional environment. Many of these policies, however, including those directed at

changes in the marginal tax rates, the share of state (vs. market) control of the economy, or the rate of innovation and technological change, all have a strong effect on investment *incentives*, and thus have the potential for an immediate impact on inequality and growth, *in the same direction*. Analogously to the analysis of restrictive institutions across countries, estimating the impact of inequality on growth within countries may be better informed by first controlling for variation in institutional incentives.

Formally, let  $R_i$  represent the level of (time-invariant) institutional restrictions to formality in country  $i$ , and let  $I_{it}$  represent a (time-varying) measure of institutionally-determined investment incentives. The rate of average income growth,  $g_{it}$ , and Gini coefficient measuring income inequality,  $G_{it}$ , are both presumed to be functions of  $R_i$  and  $I_{it}$ . For simplicity, we will ignore for the moment other potential growth and inequality determinants and assume that the period average of the incentives measure  $I_{it} = \bar{I}_i$  is normalized to be zero in each country. We believe that growth and inequality can be thought of as being determined, to a first order approximation, by the following system of equations:

$$(1) \quad g_{it} = \alpha_0 + \alpha_1 R_i + \alpha_2 I_{it} + \varepsilon_g$$

$$(2) \quad G_{it} = \beta_0 + \beta_1 R_i + \beta_2 I_{it} + \varepsilon_G$$

where our maintained hypothesis,  $\mathbf{H}$ , is that  $\alpha_1 < 0$  and  $\alpha_2, \beta_1, \beta_2 > 0$ .

First-wave regressions entail a reduced-form relationship between the long-run averages of growth on inequality across countries. Again ignoring additional controls for the moment for simplicity, this specification can be written:

$$(3) \quad \bar{g}_i = c_0 + c_1 \bar{G}_i + u_i$$

Under our set of parameter assumptions,  $\mathbf{H}$ , the coefficient estimate on inequality will converge to a negative value, or  $c_1 \xrightarrow{p} \frac{\alpha_1}{\beta_1} < 0$ . Similarly, estimates on the same coefficient derived from the fixed effects specification

$$(4) \quad g_{it} = \mu_i + c_1 G_{it} + u_{it}$$

used in the second-wave of regressions will converge to a positive number, as

$$c_1 \xrightarrow{p} \frac{\alpha_2}{\beta_2} > 0.$$

If we continue to assume for the moment that there no other variables are driving growth or inequality aside from our two measures of institutional restrictions and incentives, we also have something to say about relative sizes of these coefficients. Between countries, the ratio of the standard deviation of growth to inequality, will equal

$$\sqrt{\frac{\text{var}_i(g_{it})}{\text{var}_i(G_{it})}} = \sqrt{\frac{\alpha_1^2 \text{var}(R_i)}{\beta_1^2 \text{var}(R_i)}} = \left| \frac{\alpha_1}{\beta_1} \right|$$

or the absolute value of the probability limit of the coefficient on the cross-sectional, or “between” regression. Similarly, the ratio of standard deviations within countries will equal the plim of the coefficient from the fixed effects, or “within” estimate.

Data from the bottom of Table 2.1 on the standard deviations of growth and inequality between and within countries suggest that, were our very simplified model correct, we would expect coefficient estimates from within-country estimates around

three times larger in absolute value than between country estimates.<sup>5</sup> In the actual regressions the within-country estimates are indeed larger, although only by a factor of two.

The framework employed above to present our argument is obviously a very simple view of the underlying data generating process: the determinants of economic growth and inequality in the real world are certainly far more complicated. Nevertheless, there is a strong *prima facie* case for our interpretation of the evidence. In particular, our belief that barriers to formality may play an important role in explaining the presence of the “omitted variable bias” suggested by Forbes (2000) is leant support by the evidence shown in Figures 2.1 and 2.2. The output share of the informal sector, as reported in Friedman et al. (2000), serves as one indicator of possible barriers to formal sector participation. The size of the informal sector is both negatively correlated with growth (Figure 2.1) and positively correlated with inequality (Figure 2.2). More than that, however, this one variable alone can explain roughly one-half of the variation in Forbes’ estimated fixed effects (representing the net effect of omitted country-specific growth determinants), and one-quarter of the variation in inequality across countries.<sup>6</sup>

De Soto (1989), La Porta et al. (2002) and Bertola et al. (2003) argue the size of the informal sector may reflect underlying institutional factors, such as the regulation of entry and labor markets. Formal sector participation represents just one aspect of institutional restrictions on economic opportunity, however. Others include access to credit, land, education and political and judicial redress. Moreover, like formal sector

---

<sup>5</sup> The actual coefficient magnitudes are much smaller than the predicted magnitudes, however (-0.04 vs. -0.19 for the between-country estimate, and 0.08 vs. 0.59 for the within-country estimate).

<sup>6</sup> Although the causal interpretation of this correlation is not clear, the marginal effect of formal sector participation on both growth and inequality is strengthened when instrumented by settler mortality rates from the 1500s, an approach motivated by Acemoglu, Johnson and Robinson (2001).

participation, these measures of economic exclusion are themselves likely to have deeper underlying institutional determinants, such as respect for property rights, the rule of law, and political freedom. We recognize, of course, that these institutional determinants may be correlated with omitted country-specific determinants of the growth process, and therefore use instrumental variables techniques to control for potential endogeneity problems following the approaches taken by Hall and Jones (1999) and Acemoglu, Johnson and Robinson (2001, 2002). Figures 2.3 and 2.4 show that the impact of formal sector participation on growth and inequality, respectively, becomes even *more* pronounced when one instruments using the log of settler mortality rates in the 1500s as suggested by Acemoglu, Johnson and Robinson (2001).

Our paper pursues these arguments further as follows. In Section 3, which follows, we extend discussions in the literature of the role of institutions on growth to examine the impact on inequality. Section 4 includes institutional measures in growth cross-country regressions and finds that inequality is no longer significant. Section 5 allows for possible correlation between inequality, institutions and the omitted determinants of economic growth, using instrumental variables techniques and simultaneous equations estimation. Section 6 returns to panel estimation, using Hausman-Taylor random effects estimator to estimate the impact of time-invariant and time-varying institutions simultaneously. Section 7 concludes.

### ***3 Institutional Foundations of Inequality***

This section investigates the institutional foundations of income inequality. One of our central contentions is that the institutional variables that have been prominently linked to economic development, particularly variables that measure the protection of private property and contractual rights, are also fundamental determinants of income inequality. To put the matter somewhat strongly, what is commonly referred to as the “quality of institutions” might with equal validity be thought of as the “equality of institutions.”

Measures of the rule of law and freedom from expropriation implicitly include some notion of equality. Where the private property rights are not well protected through public institutions involving impartial courts and police, agents will use private resources to protect their property. In this case, there will be unequal protection of property reflecting the inequality of wealth and political power which agents have at their disposal. While Robert Mugabe may enjoy excellent property rights protection in contemporary Zimbabwe, this alone is insufficient to support successful development. Weak property rights invariably lead to unequal protection.

The link between institutions and income inequality is explicit in the series of papers by Sokoloff and Engermann (2000, 2002) addressing the comparative development of North and South America. The Sokoloff and Engermann suggest that resource endowments and density of the indigenous population present at the time of colonization had a direct effect on early institutional development. Where there was an abundant supply of indigenous labor and conditions were appropriate for large scale plantations and mining operations, immigration was limited and European colonizers adopted institutions that tended to concentrate economic and political power in their

hands. Over time these institutions simultaneously raised income inequality and blunted economic development: “Members of the elites were better able to maintain their elite status over time, but at the cost of society not realizing the full economic potential of disadvantaged groups” (2000, p. 228-230).

A similar link between institutions and inequality may be found in the work of Acemoglu, Johnson and Robinson (2001, 2002), henceforth AJR. AJR find that in colonies where European mortality rates were high, settlers adopted “extractive institutions” that tended to retard development. What many readers appear to have overlooked in this research is the explicitly distributional dimension of AJR’s understanding of extractive institutions, “which concentrate power in the hands of a small elite and create a high risk of expropriation for the majority of the population, are likely to discourage investment and economic development.” (2002, p. 1235) In AJR’s writing, it is not sufficient that someone has secure property rights, but rather that a reasonable degree of security is shared by the broad mass of society.

We begin by investigating the proximate determinants of inequality. Our dependent variable is the average level of income equality for each country, defined as the average of all observations for a given country between 1960 and 2000 to meet the quality standards noted above. The downside of this strategy is that, depending on data availability, we use inequality measured at very different points in time for different countries. In defense of this strategy, we note that inequality appears to be highly stable within countries over time, and that using average inequality dramatically increases the number of observations we have to work with.

Our independent variables are chosen as proximate determinants of income inequality, a number of which proxy for inequality of investment patterns and asset distribution. We use primary enrollments in 1970 (*p70*), as reported in Barro and Lee (2000), as a proxy for the inequality of human capital investment. Inequality of land ownership (*lginif*) is proxied by the Gini for farmsize, as reported in various issues of the FAO and compiled by Frankema (2005). Finally, as a proxy for the inequality of capital ownership, we use a measure of financial development – the share of private credit in GDP (*priv0*) – as reported by Beck and Levine (1999). Levine and coauthors (1999, etc.) argue that financial underdevelopment disproportionately reduces access to credit by the poor and is an important determinant of income inequality.

We also include independent variables that may influence income differentials in a manner that is independent of asset inequality. The empirical relationship between inequality and formal sector participation has been established in Section 2. Since the same resource may earn different returns if employed in the formal and informal sectors, the informal sector share of output may have a structural impact on income inequality (see also Rosser, Rosser and Ahmed 2000, 2004).

We allow for two additional sets of controls. The first is the level of per capita income in 1960 and its square, which we use to control for the Kuznets-curve pattern in the data that has been found to be an empirical regularity in a cross-section by a number of authors including Li, Squire and Zou (1998), Barro (2000), and Hopkins (2005). The second is a measure of the intensity of redistribution.<sup>7</sup> We use two variables attempt to

---

<sup>7</sup> Redistribution takes numerous forms and may affect different inequality measures differently. Redistribution in kind will affect the distribution of consumption, while taxes and transfers will affect the distribution of disposable income. Covert forms of redistribution, such as may occur through minimum wages and public employment, will affect the distribution of earnings. In addition, in sum redistributive

capture the impact of redistribution on income inequality. One variable is government consumption as reported in the Penn World Tables 6.1. Government consumption is measured as a share of domestic output and averaged over the years 1970 to 1990 (*gc7090*). This variable may reflect the extent of government employment or more broadly a taste for state intervention, including a willingness to use the state to redistribute income.

A large public sector, however, may also be predatory or evidence of redistribution from the poor to the rich or politically well-connected. For this reason, we also include a measure of democratic rights in 1970. A large body of research suggests that the intensity of progressive redistribution will rise with the political power of the poor, e.g. Bourguignon and Verdier 2000, Benabou 2000. Our measure of democratic rights is the polity variable, taken from the Polity IV dataset. The polity variable is simply a country's democracy score minus its autocracy score, each of which range from 0 to 10. The resulting index is a measure of the overall level of democratic institutions in a country, ranging from -10 (least democratic) to +10 (most democratic).

The main results are presented in Table 3.1. The first column includes the income regressors and our measures of asset inequality, of which only the land gini is significant at traditional levels. Adding the informal sector share of output, we see in column two that all of our regressors are significant with the exception of private credit. Because this variable does not appear to be significant and also limits our sample size we drop it in the remaining regressions.

---

programs may either increase or decrease income inequality. That is, redistribution may be from rich to poor or vice-versa. Since the inequality measures used are either taken from, or adjusted to match, gross household income, transfers will have a direct impact on measured inequality while progressive or regressive taxation will have only an indirect effect, if any.

Column three contains our preferred regression, including two forms of asset inequality as well as informal sector share. Each of the included regressors is significant at the 5% level or higher, and together they account for 57% of the observed cross-country variation in income inequality. In columns four and five we introduce, one at a time, government consumption and democracy in 1970, our two proxies for redistribution. Neither of these variables proves to be significant at standard levels, and the coefficients on the other regressors remain remarkably stable.

The regressions in Table 3.1 suggest that asset inequality and the types of economic marginalization that gives rise to a large informal sector are both important determinants of cross-country income inequality. Taken together, they suggest that income inequality results from economic exclusion that occurs along a number of dimensions, including the educational arena, access to land ownership and participation in the formal sector of the economy. Surprisingly, access to credit did not show up as an important determinant of income inequality.

We turn next to the institutional foundations of economic exclusion. We rely primarily on three variables to measure a country's institutional environment. For economic institutions, we use the "freedom from expropriation" variable reported in ICRG surveys. This variable has been used as a measure of property rights protection in a number of studies, including Keefer and Knack (1995) and Acemoglu, Johnson and Robinson (2001). We average expropriation risk over all available years, so that we obtain a single observation for each country. Political institutions are measured by democracy in 1970, as reported in the Polity IV dataset. Finally, we proxy legal institutions by including a dummy variable for French legal heritage. An emerging body

of research has found that a country's legal heritage is an important determinant of its current legal and regulatory environment. Relative to the English common law system, having a legal system that derives from French civil law is associated with greater state power vis-à-vis the private sector, more regulatory oversight, and a less adaptable legal framework.<sup>8</sup>

Table 3.2 provides evidence that institutional variables influence economic exclusion along a number of dimensions. It is well known that measures of institutional quality such as property rights protection are positively correlated with income level. To avoid confusing the effects of institutions with the more general impact of economic development, we include per capita income in 1970 in each of the regressions in Table 3.2.

Turning now to column one, we see that primary enrollment rates provide the least compelling evidence of a strong role for institutions. Neither economic or political institutions appear to be important. This is particularly interesting since one of the central lines of the theoretical work on inequality and growth suggests that political rights should result in greater redistribution in the provision of public education. This line of theory is not supported by our evidence. Column one does, however, suggest that a French legal heritage increases primary enrollment rates. While the main line of research on legal heritage concerns its impact on regulation, not the provision of education, this finding is consonant with the idea that French legal heritage is associated with a greater role for the state.

Turning to column two, we see that both freedom from expropriation and French legal heritage are significant determinants of financial development. Furthermore, both

---

<sup>8</sup> See Beck and Levine (1999?) for an excellent overview of this topic.

institutional variables have the expected sign, with the financial sector responding favorably to property rights protection and negatively to regulation. In column three, we add credit law as a control variable. Credit law is a measure of the protection of creditor rights used by Beck and Levine (XXX). It is not significant in this regression. More importantly, our measures of economic and legal institutions remain significant.

Column four provides evidence that both the level of development and expropriation risk are positively associated with land inequality. We believe that the positive coefficient on income captures some aspect of the commercialization of agriculture. The negative coefficient on freedom from expropriation allows two interpretations. First, it may be that in the absence of public property rights enforcement the poor are unable to defend their property from confiscation leading to a concentration of property in the hands of the rich or well-connected. Alternately, the inequality of farm size may reflect the presence of a landed elite that sees effective property rights as undermining its position of privilege. In this case, causation would flow in the opposite direction, from high land inequality to low levels of property rights protection. In either case, this connection between land inequality and property rights should not be stressed too highly, as the significance of economic institutions does not survive the introduction of geographic variables in column five. In addition, neither regression has much explanatory power, with R-squared below .33 in each case.

Columns six and seven consider the determinants of informal sector size. Our results support previous work that has stressed the roles of regulatory and red tape barriers to formality, e.g. De Soto (1989), Djankov et al. (2002), Bertola et al. (2003), Friedman et al. (2000). We find that freedom from expropriation is a highly significant

determinant of informality. This suggests that when property rights are insecure, agents may employ their assets in the informal sector in an attempt to avoid confiscation.

Alternately, one of the primary benefits of formality is access to the formal institutions, including the protection of property rights. When these fail to function well, the relative benefits of formality are significantly reduced.

In column seven, we replace the dummy for French legal heritage with a number of variables that proxy for particular aspects of a country's regulatory environment. These variables include proxies for the regulation of entry from Djankov et al. (2002) and the regulation of employment contracts and the legal strength of unions from Bertola et al. (2003). The significance of freedom from expropriation survives the introduction of these regulatory variables, none of which proves to be significant.

In general, our results in Table 3.2 support the contention that economic, legal and political institutions are important determinants of the economic exclusion that leads to income inequality. Our results are strongest for informality and financial development, and relatively weak for land inequality and primary enrollment rates.

We turn next to reduced form regressions of income inequality on institutions. These are reported in Table 3.3. Column reports the results of a regression of average inequality on our three institutional variables, our four proxies for economic exclusion, and income and income squared. With the institutional variables included in the regression, most of the measures of economic exclusion are now insignificant (compare to column three, Table 3.1). Only land inequality remains significant. This result is unsurprising since our investigation in Table 3.2 suggested that institutional variables do not explain much of the variation in land inequality. Of the institutional variables, our

proxy for economic institutions is significant at the 1% level. Political and legal institutions do not appear to be significant, but this may be because they are highly correlated with economic exclusion.

In the second column, we regress income inequality on institutions and land inequality. The other measures of economic exclusion are dropped since they do not appear to exert an independent influence on inequality. As in column one, both economic institutions and land inequality are statistically significant determinants of income inequality. The partial correlation between income inequality and freedom from expropriation is shown in Figure 3.1.

Political rights are also significant at the 5% level. Greater democratic rights are associated with lower levels of inequality. The coefficient on French legal heritage is negative, but not significant at conventional levels. One possible interpretation of this result is that French legal heritage has multiple effects that partially or fully offset each other. For example, in our earlier analysis, we found that French legal heritage is associated with a higher primary enrollment, which decreases inequality, but lower financial development and higher land inequality.

The remaining two columns test for robustness by adding regional and geographic variables to the regression in column two. Our measure of economic institutions remains significant in the presence of geographic variable but is not significant when regional dummy variables are included in the regression. Similarly, our measure of democratic rights remains significant when regional dummies are introduced, but they are insignificant when geographic variables are included.

#### ***4 Institutions should be included in Inequality-Growth Regressions***

The empirical growth literature has supplied a great deal of evidence that economic, political and legal institutions, particularly our proxy for the protection of property rights, play an important role in determining the rate of economic development across countries. We have shown in the previous section that these play an important role in determining the level of inequality across countries as well. Now we will argue that the omission of institutional variables from growth regressions has resulted in biased coefficients and a misinterpretation of the role of inequality in economic growth.

We present two forms of evidence in support of this claim. First, we examine standard first-wave style cross-country growth regressions that led to the initial claim that inequality is negatively related to growth. We demonstrate that when our institutional variables are introduced into these regressions, inequality is no longer significant (while our measure of economic institutions is). This result is robust to the inclusion of regional dummy variables and geographic variables. We conclude that the omission of economic institutions from cross-country growth regressions has led researchers to mistakenly conclude that inequality is bad for growth. Our results suggest a better interpretation of the evidence is that low quality economic institutions result in both high inequality and low economic growth.

We then move to a second-wave fixed effects estimator in which the impact of time-invariant variables is absorbed into the country-specific intercepts that capture the permanent component of a country's growth experience. We have shown that these fixed country effects are strongly associated with the degree of formality and economic

exclusion, and argued that formality is driven in part by institutional variables. Here we confirm that the country-effects are indeed largely determined by institutional variables. Freedom from expropriation alone explains over 65% of the variation in the country intercepts. Again, our conclusions are robust to the inclusion of geographic and regional dummy variables.

### **Institutions and Cross-Country Growth Regressions**

The conclusion that inequality is negatively related to economic growth reported in the first wave of growth and inequality regressions have been viewed as unreliable due to potential omitted variable bias. The first column of Table 4.1 presents a baseline growth regression, in which the dependent variable is the average growth rate of real per capita income from 1970 to 1995 with initial income included as an independent variable to control for conditional convergence along with primary enrollments in 1970, a measure of human capital investment. We do not include the investment rate because of its link to the saving rate, which many theories imply may depend on the level of income inequality. Instead we use two variables that influence investment decisions, the domestic price of investment goods, averaged from 1970 to 1990, and private credit, a measure of financial development. In our baseline regression we find that initial income, primary enrollment rate and financial development are all statically significant at the 1% level.

In column two, we add two measures of inequality, average income inequality and land inequality. Our findings are similar to those of the first-wave regressions: the effect of inequality on the long run average growth rate is negative and significant. As

mentioned previously, however, we believe that omission of institutional variables may seriously bias the coefficient on income inequality. The regression reported in column three lends credence to this claim. In this regression we have included our proxies for the quality of economic and political institutions, freedom from expropriation and democratic rights in 1970. As in columns one and two, we continue to find strong evidence that initial income, primary enrollment rates and financial development matter for economic growth. Turning to the institutional variables, we see that freedom from expropriation, our measure of property rights protection, is statistically significant at the 1% level. Freedom from expropriation is also economically significant. The regression suggests that an increase in freedom from expropriation of one standard deviation (1.83 points on a ten-point scale, or roughly the difference between Panama at 5.66 and Chile at 7.5) is associated with a 1.15 percentage point increase in the average annual growth rate. The other institutional variables do not appear to be significant determinants of economic growth. The coefficient on political institutions is not statistically significant.

More importantly for our purposes, the introduction of our institutional variables has a dramatic impact on our results for income inequality. The coefficient on income inequality falls in magnitude by nearly two-thirds. In addition, it is no longer statistically significant. These results clearly support our contention that the omission of institutional variables led to biased coefficient on income inequality in earlier cross-country regressions.

Both income inequality and institutional variables are known to vary dramatically across regions, raising the possibility that our results in column three are driven by omitted regional variables. In column four, we check this possibility by repeating the

previous regression with the introduction of regional dummy variables. This reduces the magnitude and significance of primary enrollment rates and private credit. However, the results for our primary variables of interest are largely the same. Freedom from expropriation remains significant at the 1% level, and its coefficient is greater in magnitude. This provides strong evidence that our measure of economic institutions is not simply a proxy for omitted regional variables. Indeed, the regression supports the idea that institutional differences within regions are more important than those between them. The coefficient on income inequality now slightly positive, but it remains statistically insignificant. Column four also reports the F-statistic for the joint significance of the regional dummy variables,  $F(5, 39) = 4.53$ , indicating we can reject the hypothesis that these variables are jointly insignificant at the 1% level.

Institutional quality is highly correlated with geography, with more tropical countries suffering from lower levels of property rights protection. Indeed, Hall and Jones (1999) use distance from the equator as an instrument for institutional quality. Because a number of potentially important determinants of economic growth – disease burden, fertility of the land and average temperature – are also correlated with geography, the coefficient on institutions reported in column three may be biased by the omission of these variables.

To consider this possibility, column five repeats the regression in column three with the introduction of three geographic control variables: two dummy variables, *tropical* and *landlock*, and *distequat*, the absolute distance from the equator. The results for both the magnitude and significance of the coefficients are very similar to those reported in column three. In particular, freedom from expropriation remains significant

at the 1% level, while income inequality remains insignificant and mildly negative. In addition, column five reports the F-statistic for the joint significance of the three regional variables,  $F(3, 41) = 0.78$ , implying we can reject the joint significance of the three variables at conventional significance levels,  $p = 0.51$ . (The joint insignificance of these variables is important below, where it contributes to our argument that they constitute valid instruments for institutional quality.)

### **Are Institutions the Omitted Variables In Question?**

As suggested by the evidence presented above, institutions may play a key role in determining both income inequality and economic growth. In addition, they appear to be highly persistent. North (1990) attributes institutional stability to increasing returns, broadly construed, which generates multiple stable institutional equilibria. Acemoglu, Johnson and Robinson (2001) provide empirical support for a high degree of institutional persistence and suggest a number of mechanisms that could lead to institutional persistence, including sunk costs, complementarities with existing investments, and continuity of the size or identity of local elites. Complementarities between formal institutions and highly persistent informal institutions may also contribute to institutional stability. The relative stability of institutions over time suggests they may constitute a key element of the time-invariant omitted variables for which the fixed effects estimator controls.

Using the fixed-effects specification reported in Table 2.1, we generated a vector of country fixed effects (a variable we refer to as *fe*). If this regression is correctly

specified, then the intercept captures the permanent component of a country's growth rate (that is, the portion of a country's growth that is explained by a country's permanent characteristics). The variation in these country intercepts is surprisingly large, ranging from -9.73 to 6.25 percentage points, with one standard deviation equal to 4.26.

In Table 4.2, we regress the country intercept on our measures of economic and political institutions. As reported in column one, the two institutional variables explain 66% of the variation in the country intercepts. Most of the explanatory power of these variables comes from the quality of economic institutions. Freedom from expropriation is significant at the one percent level, and alone explains 65% of the variation in the intercepts (see Figure 4.2). Our measure of political institutions is not significant in this regression. We interpret this regression as supporting the hypothesis that economic institutions are an important element of the omitted time-invariant omitted variables.

The remaining columns of Table 6 lend support to this interpretation by conducting simple robustness tests in which institutional covariates are added as regressors. As institutional development often parallels economic development, the omission of income may bias the results reported in column one. Indeed, when we include the income level as a regressor in column 2, we find that this variable is highly significant and greatly improves the explanatory power of the regression. In addition, our measure of political rights is now both negative and significant at the 5% level, though as seen below the negative association between democracy and the permanent component of growth does not appear to be robust. Finally, the inclusion of income as a regressor reduces the coefficient on freedom from expropriation by about one-third. In spite of this, however, it remains highly economically and statistically significant. Figure 4.2 shows

the partial correlation between freedom from expropriation and the permanent component of economic growth, using the specification in column (2) of Table 4.2.

As noted in section 3, institutional quality is also correlated with regional and geographic variables, and their omission may also lead to biased coefficients in column one. In column three, we add regional dummy variables to the regression in column two. In this case, political institutions are no longer significant. The quality of economic institutions, however, remains both economically important and statistically significant at the 1% level. The coefficient on freedom from expropriation is also highly stable, suggesting institutional quality is not a proxy for regional omitted variables. Regional variable are, however, significant determinants of the country intercepts. Column three also reports the F-statistic for the joint significance of the regional variables,  $F(6, 58) = 8.93$ , implying that we may reject the hypothesis that the regional coefficients are jointly zero at the 1% level.

In column four of Table 4.2, we check whether the importance of economic institutions is overstated due to the omission of geographic variables. The geographic variables are jointly significant, and we can reject the hypothesis that their coefficients are jointly zero at the 5% level. However, our proxy for the quality of economic institutions remains highly significant, suggesting that omitted geographic variables are not the reason it was found to be significant in the earlier regressions. In addition, the magnitude of the coefficient on economic institutions is very similar to that in columns two and three. Finally, our measure of democratic rights is now significant at the 10% level and negatively related to the permanent component of a country's growth rate.

The evidence presented supports the hypothesis that economic institutions are an important determinant of the permanent component of a country's growth rate. Freedom from expropriation alone explains about 2/3 of the variation in the country intercepts, and the importance of this variable does not appear to be a statistical artifact derived from the correlation between the quality of economic institutions and the level of economic development, omitted regional variable or geography. In short, it is only a minor stretch to suggest that economic institutions *are* the time-invariant omitted variables that fixed effects growth regressions were designed to avoid.

## 5 Discussion of Collinearity, Endogeneity and Plausible Instrumenting Procedures

This paper, like others in the now-substantial empirical growth literature, has attempted to gain insight into the determinants of growth by revealing the conditional relationships that exist between national growth rates and various economic characteristics. Caution is required when drawing inference from cross-country regressions, however. The reasons are, for the most part, widely-recognized and have been discussed in detail elsewhere (e.g. Mankiw, 1995; Brock & Durlauf, 2002). For this reason, we will focus our discussion on only one of these, although possibly the most important.

The wide variation in development experiences across countries provides a rich terrain for estimating the impact of economic conditions on economic growth. This very heterogeneity of experience, combined with the vast number of confounding factors determining growth rates, makes statistical inference difficult. Appropriate statistical inference requires that residuals from the growth regression represent repeated,

independent draws from the same distribution. Given the manifest differences among countries such as Tanzania, France, Korea and India, the *i.i.d.* assumption can only be considered valid after an appropriate conditioning process. This conditioning process is likely to be flawed for [three] reasons, each of which are essentially problems of small sample size. First, the correct specification may require a high degree of parameterization, which presents a degrees-of-freedom problem in a limited cross-country sample.<sup>9</sup> Secondly, arriving at such a specification in the first place is very difficult because of strong collinearity among variables, which makes it difficult to carefully discriminate among competing specifications. Finally, in the cross-country setting, many of the omitted determinants of observed growth rates are very likely to be correlated with variables appearing on the right hand side of the regression equation, violating statistical exogeneity assumptions.

To deal with these issues, we – like much of the literature – have employed a combination of three approaches:

- 1) A heavy reliance on *a priori* reasoning to justify assumptions made regarding specification and exogeneity,
- 2) parsimonious regressions aimed at uncovering specific reduced form relationships among a few variables, rather than arguments relying heavily on structural interpretations of parameter estimates.
- 3) plausible exclusion restrictions to justify independence assumptions regarding certain instruments and the residuals in the growth regression.

---

<sup>9</sup> This problem is made more severe by the fact that in practice, as the rank of the design matrix increases, the available sample size tends to decrease because of missing data for many countries.

### ***Evidence with Endogenous Inequality***

The two previous sections have argued that institutions are an important determinant of income inequality and that their omission from previous growth regressions has resulted in misspecification and led to biased estimates of the impact of inequality on growth. Simply adding institutions to the cross-sectional regressions does not entirely solve the problem, however. As noted in Section 2, however, it is more appropriate to consider a structural model in which both growth and inequality are treated as endogenous.

We begin by extending the simultaneous equations framework laid out in Section 2 to include inequality as a regressor in the growth equation along with a set of additional controls. Identification will be achieved via inclusion of control variables that can reasonably be expected to effect one aspect of the income distribution (either growth or inequality) without necessarily affecting the other. Our reduced form equations are thus:

$$(5) \quad \begin{aligned} \text{growth} &= F(X, \text{inequality}, \text{institutions}) \\ \text{inequality} &= G(Y, \text{institutions}) \end{aligned}$$

which are applied first to our cross-country data set, representing first wave regressions.

In the growth equation, our control variables include initial income, the primary enrollment rate in 1970, and the average price of investment goods. These variables control for conditional convergence, human capital accumulation and physical capital accumulation. Our list of control variables for the inequality equation includes the income level in 1970, its square, the primary enrollment rate, the land Gini and the informal sector share. These control variables are the same as those used in the single-equation regressions for growth and inequality reported in Tables 4 and 1, respectively,

with the exception that we have dropped private credit from the growth regression because it restricts our sample size.

Column one of Table 5.1 reports estimates for (1) without the institutional variables. Our findings for the growth equation are similar to those reported in column 2 of Table 4. Both initial income and primary enrollment rate are statistically significant determinants of growth and have the expected signs. In addition, income inequality is negatively related to growth and statistically significant. Similarly, both the land Gini and informal sector share have the expected sign and are statistically significant.

In the second column of Table 5.1, we include our proxies for economic and political institutions as regressors. Given our earlier investigations, the impact of this change is much as we expect. With the inclusion of our institutions variables, inequality is no longer significant in the growth equation. In addition, the protection of property rights is statistically significant in both regressions. Better property rights protection simultaneously increases the rate of growth and decreases income inequality. Democratic political rights appear to involve a growth-equity trade-off, lowering income inequality but at the expense of lower economic growth, though it is not statistically significant in the inequality regression. In addition, the inclusion of institutional variables significantly improves the fit of each equation.

Our last regression is subject to a number of criticisms. The institutional variables are derived from expert opinion and survey data, and thus potentially subject to systematic measurement error. For example, if experts “see” better institutions in countries that experience higher growth rates or have lower income inequality, then this will lead to biased estimates of the institutional coefficients. In addition, the quality of

economic and political institutions typically rise with income, making these variables endogenous. And, finally, the omission of variables that are simultaneously correlated with institutions and either growth or inequality may bias our coefficient estimates.

We address these issues by instrumenting for freedom from expropriation and democracy in 1970. A large number of instruments for contemporary institutions have been used in the literature. We review these briefly below. Acemoglu, Johnson and Robinson (2001, 2002) find that settler mortality rates and population density in 1500 are important for later institutional development. The importance of pre-colonial population density is also stressed by Sokoloff and Engermann (2000), who suggest that the availability of an indigenous workforce limited European immigration and led to the adoption of institutions that served an elite European minority. Given the constraints on our sample size posed by the availability of our other variables, we do not rely overly on these variables, though they are included to check the robustness of our results.

AJR (2001) also find that early democratic politics is an important determinant of future institutional development. Again, their preferred variable, democracy in 1900, proves too restrictive given our sample, but we do include a more widely available measure, democracy in 1960, as an instrument.

A second set of instruments reflect geography. Hall and Jones (1999) argue that latitude serves as a proxy for institutional transfer during colonization, perhaps because lateral migrations allowed European settlers to transfer existing agricultural technologies. Our geographical instruments include the absolute value of latitude and its square. The use of the square of latitude allows for a non-linear relationship between latitude and institutional transfer, and significantly improves the fit of the first-stage regressions. Hall

and Jones also suggest using the fractions of a countries population that speaks English or a European language as a first language as instruments for institutional quality. We do not find that the inclusion of these variables as instruments alters our results.

A third set of instruments involves a countries legal heritage. An important emerging literature has linked legal heritage to measures of governance (Djankov et al. 1999), the efficiency of the courts (La Porta “Courts” ), and the regulation of finance, entry and labor markets (see XXXX). We include as instruments dummy variables for English, French, German, Scandinavian and Socialist legal heritage.

In the third column in Table 5.1, we treat freedom from expropriation and democracy in 1970 as endogenous. We instrument with democracy in 1960, geographic variables and legal heritage. The results reported in column 3 largely confirm our earlier findings. In particular, inequality is not significant in the growth equation, whereas the coefficient on freedom from expropriation enters significantly in both equations at the 5% level or better, and with the same signs as reported in column 2. That is, the protection of property rights increases growth and decreases inequality. In addition, the magnitude of the coefficient has increased in each equation. This suggests that measurement error – particularly the tendency to see good institutions in countries with good outcomes – may have biased the institutional coefficients in column two downward. Also as in column two, democratic political rights again appear to be negatively related to growth and unrelated to inequality.

Columns four, five and six consider robustness of these results. In column four, we include regional dummy variables for sub-Saharan Africa and Latin America and the Caribbean, both of which prove significant in at least one equation. Most of our results

are robust to the inclusion of regional dummies, although freedom from expropriation is no longer significant in the inequality regression. In column five, we add the two language variables suggested by Hall and Jones (1999) as instruments. Here, the results for our institutional variables are very much the same as they were in column 3, both in magnitude and in significance. Finally, in column 6, we replace our geographic instruments with the instruments suggested by AJR: the log of European settler mortality rate and the population density in 1500. With the use of these instruments, freedom from expropriation is no longer significant in the inequality equation, though this may simply reflect the restriction of our sample to 29 countries.

## ***6 Efficient GLS Estimation of the Growth-Inequality Relationship***

## ***7 Discussion and Conclusions***

## References

- Acemoglu, Daron, Simon Johnson, and James Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, December 2001, v. 91, iss. 5, pp. 1369-1401
- Acemoglu, Daron, Simon Johnson, and James Robinson, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, November 2002, v. 117, iss. 4, pp. 1231-94
- Alesina, Alberto and Dani Rodrik, "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*, May 1994, v. 109, iss. 2, pp. 465-90
- "Banerjee, Abhijit V. and Esther Duflo, "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth*, September 2003, v. 8, iss. 3, pp. 267-99
- Barro, Robert J. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics*, May 1991, v. 106, iss. 2, pp. 407-43
- Barro, Robert J. "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth*, March 2000, v. 5, iss. 1, pp. 5-32
- Barro, Robert J. and Jong-What Lee, "International Data on Educational Attainment Updates and Implications." National Bureau of Economic Research, Inc, NBER Working Papers: 7911. 2000
- Benabou, Roland "Inequality and Growth." *NBER macroeconomics annual 1996*, 1996, pp. 11-74.
- Benhabib, J, M. Spiegel, "Cross-Country Growth Regressions." C.V. Starr Center for Applied Economics, New York University, Working Papers., 1997.
- Bertola, Guiseppe et. al. (2003)
- Birdsall, Nancy and Juan Luis Londono, "Asset Inequality Matters: An Assessment of the World Bank's Approach to Poverty Reduction," *American Economic Review*, v. 87, No. 2. May, 1997.
- Bourguignon, Francois and Thierry Verdier, "Oligarchy, Democracy, Inequality and Growth." *Journal of Development Economics*, August 2000, v. 62, iss. 2, pp. 285-313
- Brock, William A. and Steven N. Durlauf, "Growth Empirics and Reality." *World Bank Economic Review*, 2001, v. 15, iss. 2, pp. 229-72.
- Davis, Lewis S. "Growth and Inequality Theory," mimeo, 2006.

Deininger, Klaus and Lyn Squire, "A New Data Set Measuring Income Inequality." *World Bank Economic Review*, September 1996, v. 10, iss. 3, pp. 565-91

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes and Andrei Shleifer. "The Regulation of Entry." *Quarterly Journal of Economics*, February 2002, v. 117, iss. 1, pp. 1-37

Engerman, Stanley L. and Kenneth L. Sokoloff, "Factor Endowments, Inequality, and Paths of Development among New World Economies," *Economia: Journal of the Latin American and Caribbean Economic Association*, Fall 2002, v. 3, iss. 1, pp. 41-88

Forbes, Kristin J. "A Reassessment of the Relationship between Inequality and Growth." *American Economic Review*, September 2000, v. 90, iss. 4, pp. 869-87

Frankema, Ewout; Smits, Jan-Pieter. Exploring the Historical Roots of Eastern Asia's Post-war Catch-Up Growth: A Trade Perspective, 1906-1999. *Journal of the Asia Pacific Economy*, May 2005, v. 10, iss. 2, pp. 178-94 ??

Friedman (2000)

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, January 1993, v. 60, iss. 1, pp. 35-52

Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, February 1999, v. 114, iss. 1, pp. 83-116

Hausman, Jerry A.; and William E. Taylor, "Panel Data and Unobservable Individual Effects." *Journal of Econometrics*, May 1981, v. 16, iss. 1, pp. 155

Hopkins, Mark, "The Determinants of Income Inequality: A Bayesian Approach," mimeo, August, 2005.

Knack, Stephen and Philip Keefer, "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics*, November 1995, v. 7, iss. 3, pp. 207-27

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, "Government Ownership of Banks." *Journal of Finance*, February 2002, v. 57, iss. 1, pp. 265-301

Levine, Ross, Thorsten Beck and Norman Loayza, "Finance and the Sources of Growth," The World Bank, Policy Research Working Paper Series: 2057. 1999

Li, Hongyi, Lyn Squire and Heng-fu Zou, "Explaining International and Intertemporal Variations in Income Inequality," *Economic Journal*, January 1998, v. 108, iss. 446, pp. 26-43

Loury, Glenn C. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica*, June 1981, v. 49, iss. 4, pp. 843-67

Lundberg, Mattias and Lyn Squire, "The Simultaneous Evolution of Growth and Inequality," *Economic Journal*, April 2003, v. 113, iss. 487, pp. 326-44

Mankiw, N. Gregory. "The Growth of Nations." *Brookings Papers on Economic Activity*, 1995, v. 0, iss. 1, pp. 275-310

North, Douglass C. *Institutions, Institutional Change and Economic Performance*. New York, NY. Cambridge University Press. 1990.

Partridge, Mark D, "Is Inequality Harmful for Growth? A Comment," *American Economic Review* (1998)

Perotti, Roberto, "Growth, Income Distribution, and Democracy: What the Data Say" *Journal of Economic Growth*, June 1996, v. 1, iss. 2, pp. 149-87

Persson, Torsten and Guido Tabellini, "Is Inequality Harmful for Growth?" *American Economic Review*, June 1994, v. 84, iss. 3, pp. 600-621

Rosser, J. Barkley, Jr., Marina V Rosser and Ehsan Ahmed, "Income Inequality and the Informal Economy in Transition Economies." *Journal of Comparative Economics*, March 2000, v. 28, iss. 1, pp. 156-71

de Soto, Hernando. *The Other Path*. New York: Harper & Row, 1989.

World Income Inequality Database (WIID), version 2.0a (2005).  
<http://www.wider.unu.edu/wiid/wiid.htm>

## IFIG – Tables and Figures

**Table 2.1 Results of panel data estimation**

Dependent variable = growth rate over following 5-year period

	(1) <u>Within (FE)</u>	(2) <u>Between</u>	(3) <u>GLS (RE)</u>
log GDP/capita	-3.787 (0.492) ***	-0.035 (0.332)	-0.554 (0.278) **
Secondary Years	0.884 (0.276) ***	0.411 (0.317)	0.252 (0.217)
Price of Investment	-0.007 (0.004) *	-0.021 (0.008) **	-0.013 (0.004) ***
Income Gini	0.081 (4.092) ***	-0.041 (0.022) *	-0.029 (0.019)
R-squared...			
overall	0.01	0.08	
within	0.20	0.00	
between	0.05	0.16	
N=	448		
country groups =	87		
T =	1 min		
	5.1 avg		
	8 max		

\*\*\* = significant at the 1% level, \*\* = 5%, \* = 10%

	Mean	Std. Dev.	Min	Max	Observations
<b>Growth next period</b>					
overall	2.03	2.63	-7.36	10.35	N = 448
between		1.85	-2.63	6.22	n = 87
within		1.97	-3.63	9.12	T = 5.15
<b>Log(GDP/capita)</b>					
overall	8.54	0.96	6.51	10.25	N = 448
between		0.98	6.51	10.11	n = 87
within		0.27	7.51	9.54	T = 5.15
<b>Price of Investment</b>					
overall	75.49	33.67	21.52	369.90	N = 448
between		23.60	26.95	137.56	n = 87
within		24.54	-20.86	317.46	T = 5.15
<b>Secondary Years</b>					
overall	1.47	1.16	0.01	5.09	N = 448
between		1.07	0.05	4.65	n = 87
within		0.49	-0.19	3.19	T = 5.15
<b>Income Gini</b>					
overall	43.02	9.73	22.63	70.70	N = 448
between		9.98	23.98	67.07	n = 87
within		3.35	32.30	58.18	T = 5.15

**Figure 2.1: Omitted growth variables vs. size of the informal sector ( $R^2 = 0.49$ )**

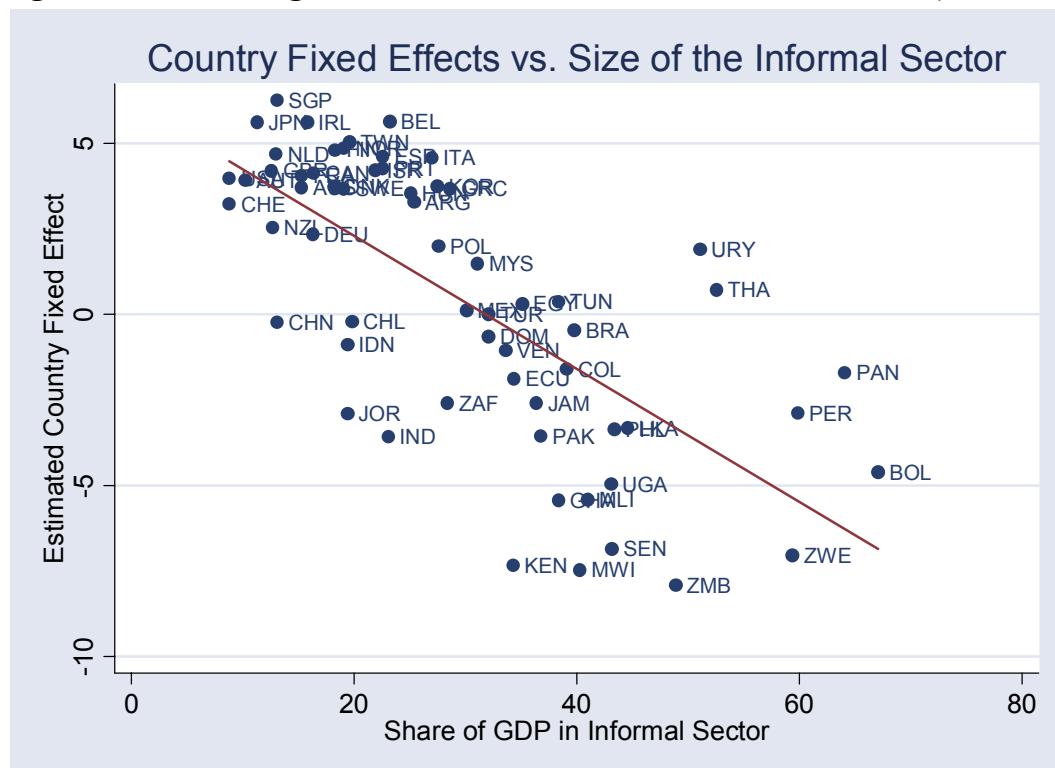
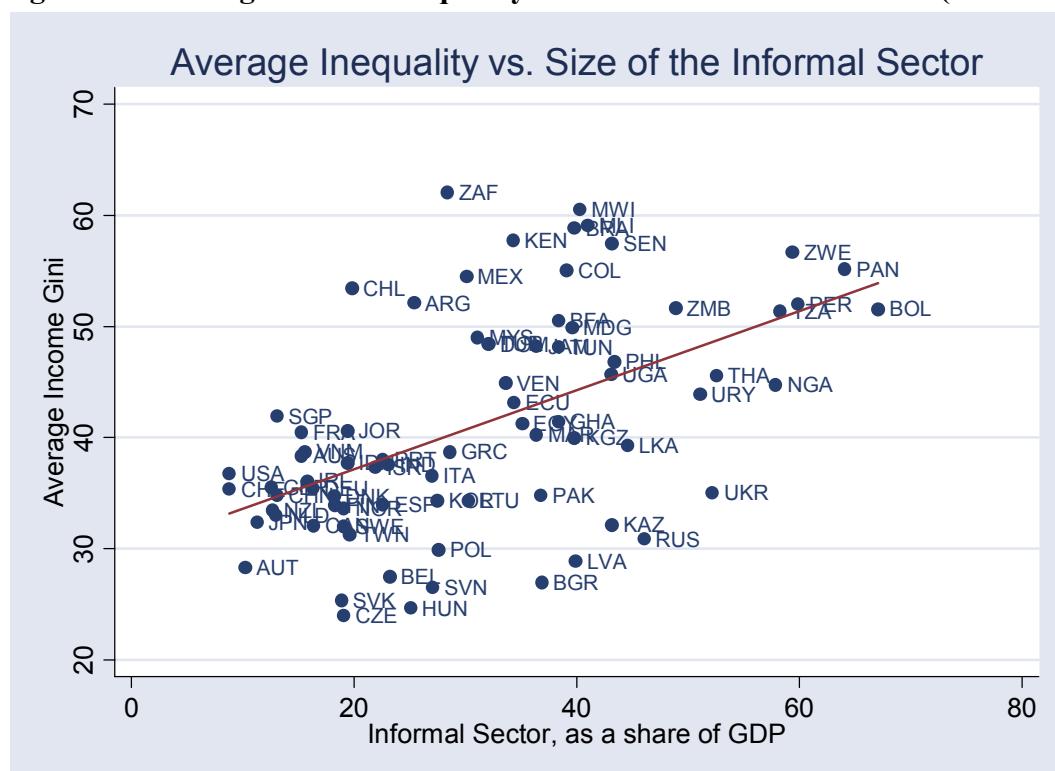
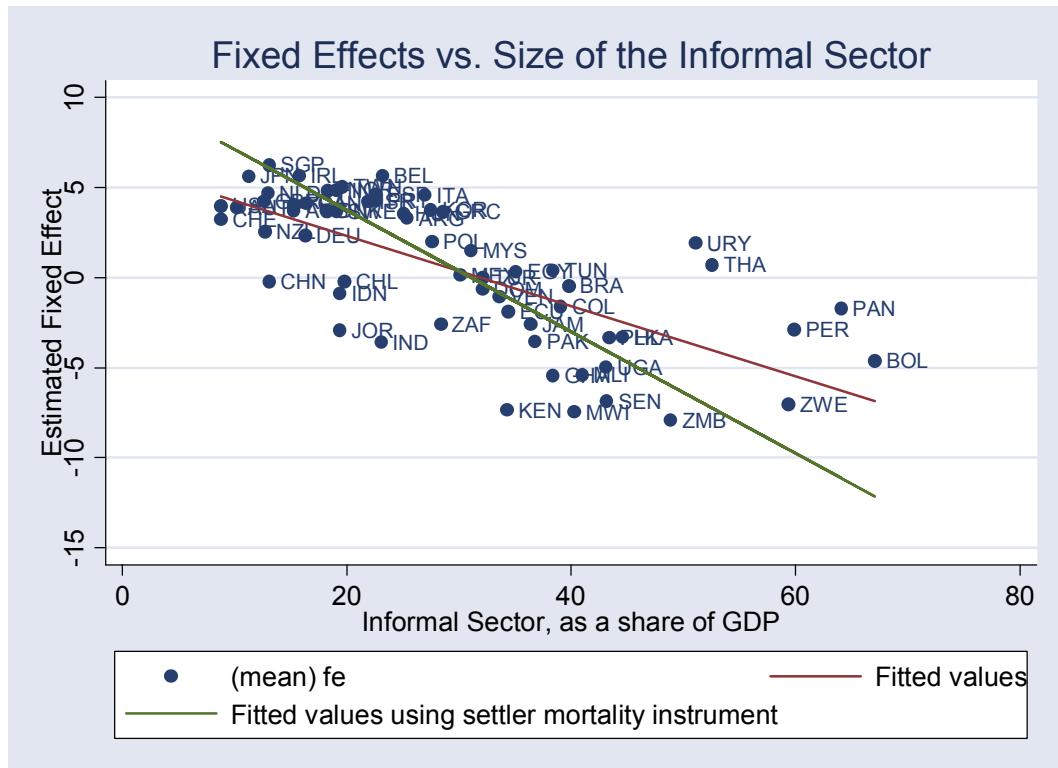


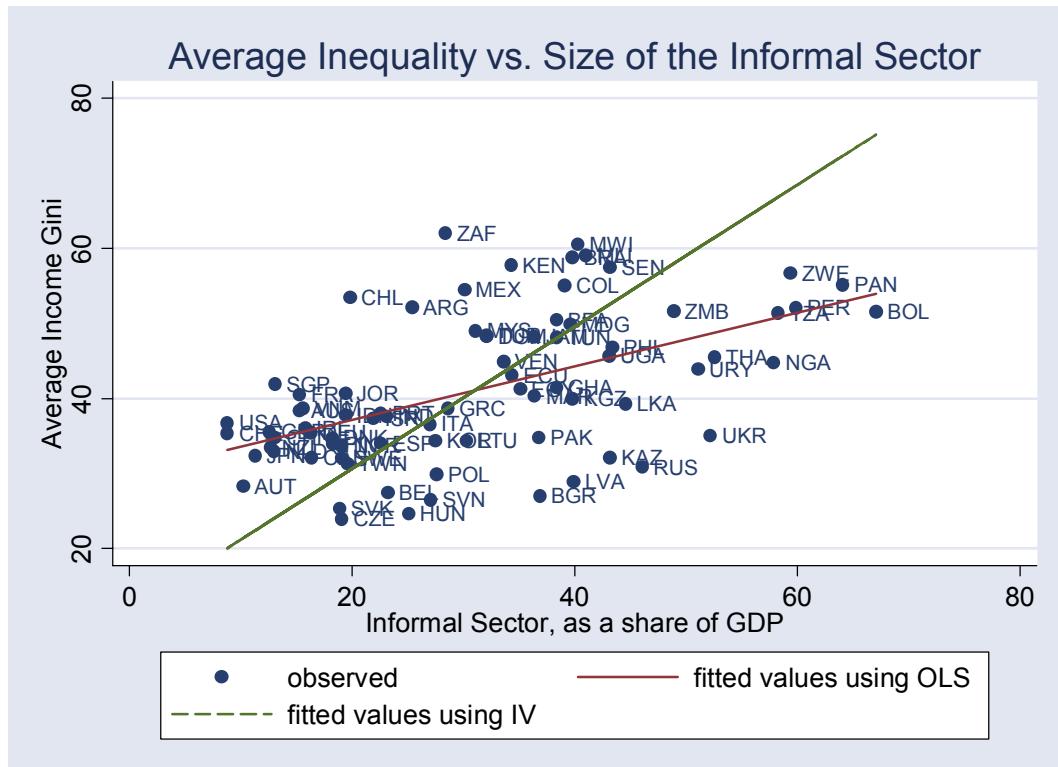
Figure 2.2 Average Income Inequality vs. size of the informal sector ( $R^2 = 0.28$ )



**Figure 2.3: Omitted growth variables vs. size of the informal sector, OLS vs. IV**



**Figure 2.4 Average Income Inequality vs. size of the informal sector, OLS vs. IV**



**Table 3.1: Inequality and Economic Exclusion**

*dependent variable = Income Gini, averaged over all periods*

	(1)	(2)	(3)	(4)	(5)	(6)
<b>lny60</b>	28.651 (23.952)	52.237* (28.792)	47.140** (22.435)	46.919** (22.604)	43.145* (23.169)	66.922** (26.168)
<b>lny60sqr</b>	-2.090 (1.473)	-3.332* (1.766)	-3.033** (1.388)	-3.029** (1.398)	-2.754* (1.444)	-4.211** (1.600)
<b>p70</b>	-0.057 (0.053)	-0.163** (0.072)	-0.122** (0.051)	-0.125** (0.051)	-0.112** (0.053)	-0.124** (0.054)
<b>lginif</b>	24.321*** (7.323)	27.902*** (7.027)	19.118*** (5.907)	19.972*** (6.160)	17.930*** (6.145)	18.080*** (5.936)
<b>priv0</b>	-0.505 (4.119)	3.903 (3.965)	--	--	--	--
<b>informal</b>	--	0.242** (0.096)	0.220*** (0.082)	0.219** (0.083)	0.223*** (0.082)	0.212** (0.085)
<b>gc7090</b>	--	--	--	-0.054 (1.000)	--	--
<b>polity1970</b>	--	--	--	--	-0.111 (0.149)	--
<b>ethfrac2</b>	--	--	--	--	--	0.053 (0.036)
N	58	45	55	55	55	54
R-squared	0.48	0.61	0.57	0.57	0.58	0.59
Adjusted R-Squared	0.43	0.55	0.53	0.52	0.52	0.54

**Table 3.2: Determinants of Economic Exclusion**

Dependent variable is....	(1) p70	(2) priv0	(3) priv0	(4) lginif	(5) lginif	(6) informal	(7) informal
<b>lny70</b>	11.390** (4.327)	0.104* (0.056)	0.054 (0.081)	0.084** (0.035)	0.071** (0.034)	-0.554 (2.261)	-1.547 (2.045)
<b>exprop</b>	1.371 (2.335)	0.086*** (0.029)	0.105** (0.044)	-0.046** (0.019)	-0.028 (0.020)	-7.387*** (1.300)	-6.671*** (1.145)
<b>polity1970</b>	0.629 (0.437)	-0.005 (0.006)	0.003 (0.007)	0.003 (0.004)	0.004 (0.003)	-0.001 (0.234)	0.243 (0.222)
<b>french</b>	13.084** (5.899)	-0.137* (0.075)	--	0.064 (0.048)	0.081* (0.046)	-2.188 (3.091)	--
<b>credit_law</b>	--	--	-0.003 (0.035)	--	--	--	--
<b>landlock</b>	--	--	--	--	-0.010 (0.063)	--	--
<b>latitude</b>	--	--	--	--	-0.003*** (0.001)	--	--
<b>tropical</b>	--	--	--	--	-0.030 (0.048)	--	--
<b>entryreg</b>	--	--	--	--	--	--	1.732 (1.635)
<b>employreg</b>	--	--	--	--	--	--	4.894 (3.724)
<b>unionreg</b>	--	--	--	--	--	--	0.426 (2.720)
N	66	57	41	69	69	58	59
R-squared	0.41	0.57	0.46	0.18	0.32	0.67	0.69
Adjusted R-Squared	0.37	0.53	0.40	0.13	0.24	0.64	0.65

\*\*\* p<0.01; \*\* p<0.05; \* p<0.1;

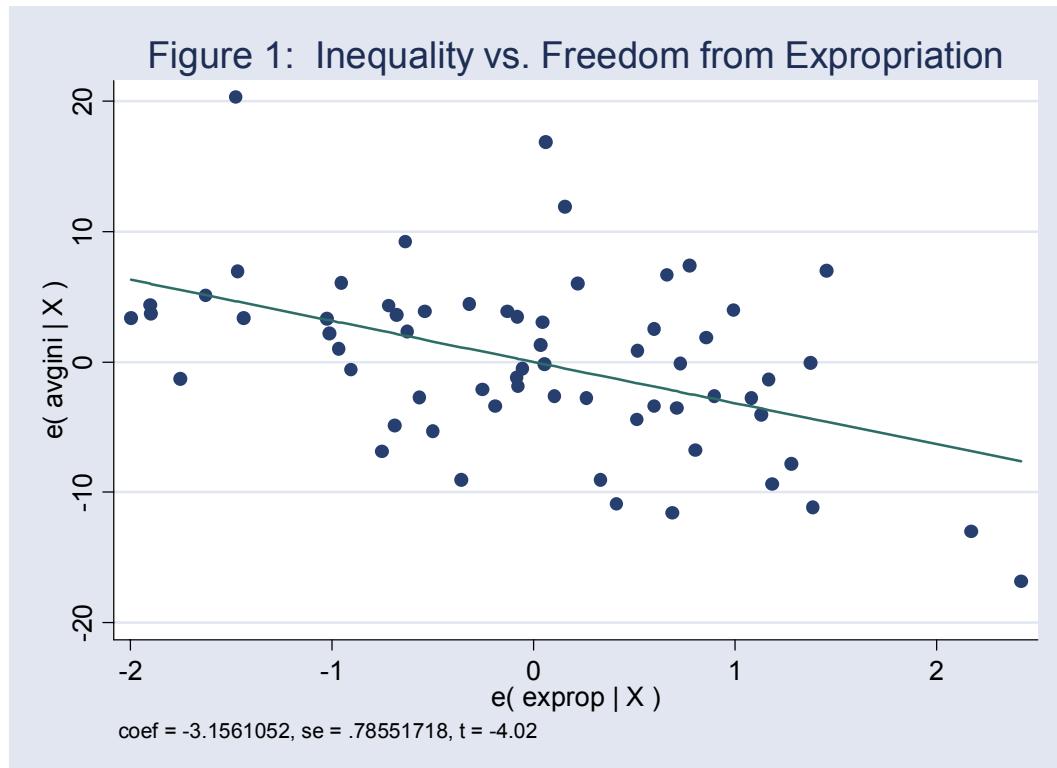
**Table 3.3 The Determinants of Inequality***Dependent variable: Income Gini coefficient, averaged over all periods*

	(1)	(2)	(3)	(4)
<b>lny60</b>	47.482 (28.300)	0.977 (19.576)	2.602 (16.301)	14.464 (18.433)
<b>lny60sqr</b>	-2.894 (1.740)	-0.055 (1.221)	-0.155 (1.017)	-0.851 (1.152)
<b>exprop</b>	-3.598*** (1.221)	-3.156*** (0.786)	-0.547 (0.815)	-1.591** (0.787)
<b>polity1970</b>	-0.168 (0.171)	-0.323** (0.154)	-0.331** (0.143)	-0.216 (0.141)
<b>french</b>	-2.360 (2.325)	-2.280 (2.181)	-2.008 (1.803)	-0.692 (1.971)
<b>p70</b>	-0.087 (0.071)	--	--	--
<b>priv0</b>	6.372 (3.866)	--	--	--
<b>lginif</b>	20.527*** (6.974)	11.001** (5.431)	6.166 (5.492)	4.938 (5.219)
<b>informal</b>	0.041 (0.111)	--	--	--
<b>reg_eap</b>	--	--	4.227* (2.474)	--
<b>reg_eca</b>	--	--	14.756** (5.575)	--
<b>reg_mena</b>	--	--	0.225 (3.252)	--
<b>reg_sa</b>	--	--	3.281 (4.635)	--
<b>reg_ssa</b>	--	--	13.559*** (3.308)	--
<b>reg_lac</b>	--	--	12.463*** (3.334)	--
<b>tropical</b>	--	--	--	4.928** (1.952)
<b>latitude</b>	--	--	--	-0.102*** (0.034)
<b>landlock</b>	--	--	--	2.610 (2.572)
N	45	64	64	64
R-squared	0.70	0.57	0.76	0.68
Adjusted R-Squared	0.63	0.52	0.71	0.63

\*\*\* p&lt;0.01; \*\* p&lt;0.05; \* p&lt;0.1;



**Figure 3.1: Partial Correlation of Average Income Gini on Expropriation Risk**



The slope in this picture represents the coefficient on expropriation risk from Regression (2) in Table 3.3.

**Figure 4.1 The Relationship of Country-specific Growth Effects to Expropriation Risk**

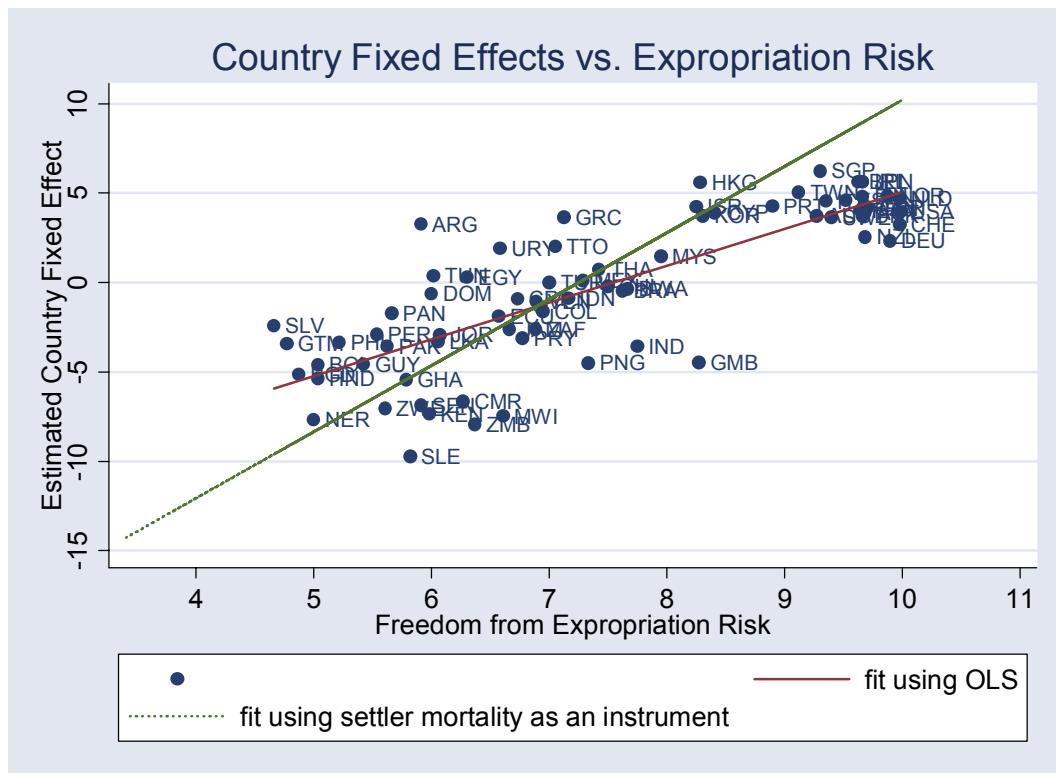
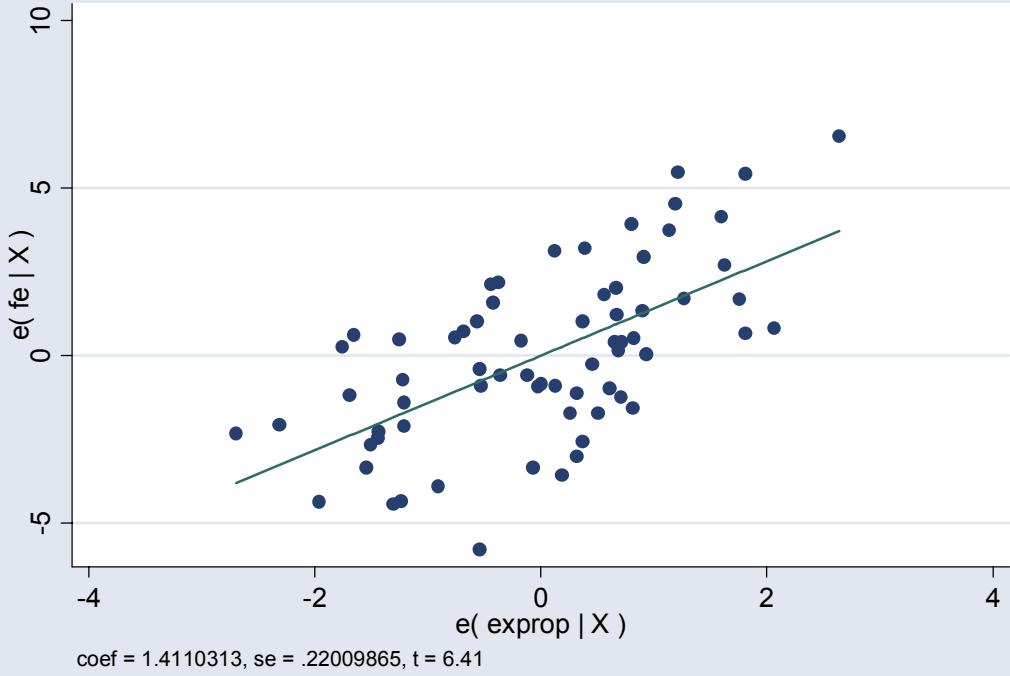


Figure 2: Country Intercepts vs. Freedom from Expropriation



*Slope represents the effect of expropriation risk, after controlling for income and democracy.*

**Table 4.1: Cross-Country Growth Regressions**

dependent variable = growth rate, 1970 - 1995

	(1)	(2)	(3)	(4)	(5)
<b>lny70</b>	-1.082*** (0.320)	-1.635*** (0.319)	-1.846*** (0.291)	-1.950*** (0.295)	-1.987*** (0.359)
<b>priv0</b>	3.199*** (0.768)	2.528*** (0.750)	1.602** (0.701)	1.148* (0.622)	1.832** (0.762)
<b>pi7090</b>	-0.010 (0.008)	-0.008 (0.008)	-0.004 (0.008)	-0.004 (0.007)	-0.005 (0.008)
<b>p70</b>	0.040*** (0.010)	0.046*** (0.010)	0.042*** (0.010)	0.022** (0.010)	0.039*** (0.010)
<b>avgini</b>	--	-0.092*** (0.023)	-0.031 (0.027)	0.016 (0.028)	-0.031 (0.027)
<b>lginif</b>	--	-0.321 (1.356)	-0.522 (1.316)	-0.739 (1.226)	0.095 (1.574)
<b>exprop</b>	--	--	0.648*** (0.182)	0.715*** (0.186)	0.557*** (0.199)
<b>polity1970</b>	--	--	-0.034 (0.025)	-0.010 (0.025)	-0.031 (0.026)
<b>Regions</b>			<i>F</i> (5,39) = 4.53		
			<i>prob &gt; F</i> = 0.002		
<b>Geography</b>				<i>F</i> (3,41) = 0.78	
				<i>prob &gt; F</i> = 0.509	
N	63	59	53	53	53
R-squared	0.38	0.55	0.67	0.79	0.69
Adjusted R-squared	0.34	0.50	0.61	0.72	0.61

\*\*\* p&lt;0.01; \*\* p&lt;0.05; \* p&lt;0.1; two tailed

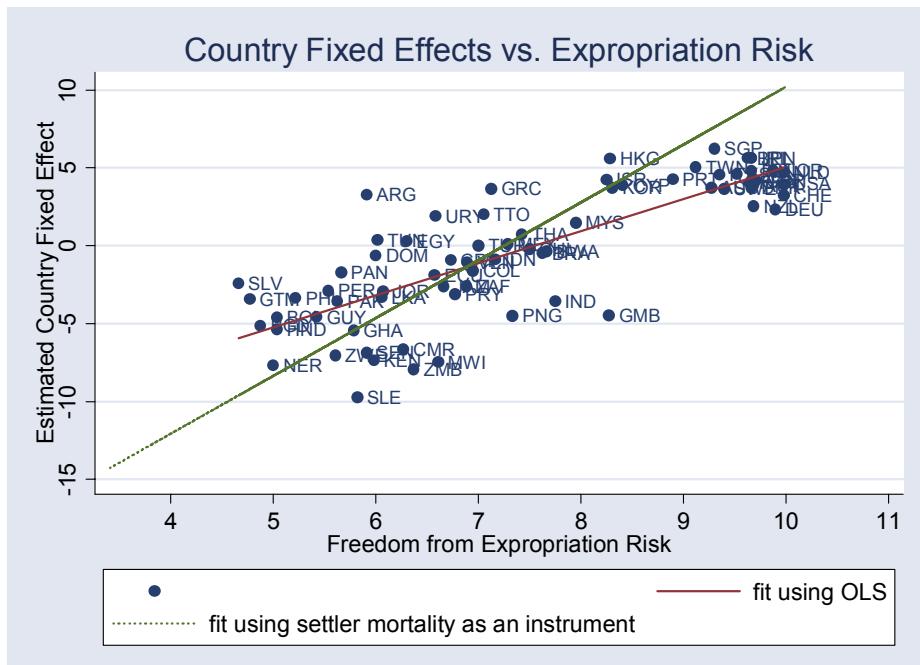
**Table 4.2: Country Intercepts and Institutions**

*Dependent variable = Estimated Country Fixed Effect from Growth Regression*

	(1)	(2)	(3)	(4)	(5)
<b>lny70</b>				1.953***	--
		(0.398)	(0.418)	(0.434)	
<b>exprop</b>	2.136***	1.411***	1.379***	1.222***	2.062***
	(0.220)	(0.220)	(0.243)	(0.242)	(0.180)
<b>polity1970</b>	-0.036	-0.094**	-0.028	-0.071*	--
	(0.051)	(0.043)	(0.037)	(0.043)	
<b>Regions</b>				<i>F</i> = 8.93	
				<i>prob(F)</i> = 0	
<b>Geography</b>				<i>F</i> = 2.98	
				<i>prob(F)</i> = 0.038	
N	68	68	68	67	72
R-squared	0.66	0.77	0.88	0.81	0.65
Adjusted R-Squared	0.65	0.76	0.86	0.79	0.65

\*\*\* p<0.01; \*\* p<0.05; \* p<0.1; two tailed

**Figure 4.1 Correlation of Country Fixed Effects to Expropriation Risk ( $r = 0.85$ , N=72)**



**Table 5.1: Simultaneous Equations Regressions using Cross-Country Data**

	(1)		(2)		(3)		(4)		(5)		(6)	
	growth	Avgini	growth	avgini	growth	avgini	growth	avgini	growth	avgini	growth	Avgini
<b>lny70</b>	-1.258	8.446	-1.391	-0.860	-2.107	89.152	-1.854	74.406	-1.938	83.219	-1.988	116.967
	(3.69)***	(0.46)	(4.21)***	(0.04)	(5.45)***	(3.23)***	(5.58)***	(3.07)***	(5.66)***	(2.96)***	(5.07)***	(3.26)***
<b>pi7090</b>	-0.004		-0.023		-0.012		-0.007		-0.014		-0.002	
	(1.55)		(2.82)***		(1.57)		(1.04)		(1.99)**		(0.22)	
<b>p70</b>	0.026	-0.066	0.019	0.007	0.017	-0.035	0.012	-0.005	0.019	-0.044	0.035	-0.035
	(2.24)**	(1.26)	(1.57)	(0.12)	(1.62)	(0.58)	(1.23)	(0.10)	(1.89)*	(0.73)	(2.51)**	(0.51)
<b>avgini</b>	-0.178		-0.108		0.058		0.058		0.010		0.046	
	(4.03)***		(1.30)		(1.07)		(1.05)		(0.24)		(1.00)	
<b>lny70sqr</b>		-0.632		0.019		-5.243		-4.493		-4.944		-7.045
		(0.57)		(0.01)		(3.20)***		(3.13)***		(2.97)***		(3.29)***
<b>LGiniF</b>	19.170		15.582		4.261		3.218		8.447		7.083	
		(3.72)***		(2.84)***		(0.75)		(0.57)		(1.43)		(0.98)
<b>informal</b>	0.263		0.110		-0.023		-0.001		0.015		0.048	
		(3.58)***		(1.22)		(0.18)		(0.01)		(0.12)		(0.31)
<b>exprop</b>		0.690	-2.192	1.601	-3.885	1.214	-1.153	1.326	-2.948	1.644	-0.194	
		(1.90)*	(2.17)**	(4.54)***	(2.16)**	(4.36)***	(0.60)	(4.70)***	(1.67)*	(5.99)***	(0.09)	
<b>polity1970</b>		-0.062	-0.105	-0.067	0.147	-0.057	0.034	-0.066	0.125	-0.137	-0.092	
		(2.19)**	(0.79)	(2.11)**	(0.81)	(2.08)**	(0.21)	(2.20)**	(0.70)	(2.97)***	(0.38)	
<b>reg_ssa</b>						-2.350	10.713					
						(2.79)***	(3.11)***					
<b>reg_lac</b>						-0.904	6.840					
						(1.50)	(2.65)***					
N	56	56	52	52	50	50	50	50	50	50	29	29
R-squared	0.19	0.55	0.54	0.60	0.58	0.60	0.73	0.73	0.65	0.62	0.60	0.55
Endogenous Vars	growth, inequality		growth, inequality		Growth, inequality, exprop, polity1970		Growth, inequality, exprop, polity1970		Growth, inequality, exprop, polity1970		Growth, inequality, exprop, polity1970	
Excluded Instruments					polity1960, legal heritage dummies, latitude and latitude squared		polity1960, legal heritage dummies, latitude and latitude squared		polity1960, legal heritage dummies, latitude, latitude squared, engfrac and eurfrac		polity1960, legal heritage dummies, augmort and density	

Notes to Table 5.1: Absolute value of t statistics in parentheses. T statistics are corrected for small sample bias. Model estimated using three-stage least squares. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

**Table 6.1 Decomposing the Inequality Coefficient Across vs. Within Countries**

	(1)	(2)	(3)	(4)	(5)	(6)
<b>log(GDP/capita)</b>	-3.787 *** (0.492)	-0.554 ** (0.278)	-0.052 (0.202)	-0.128 (0.199)	-0.660 *** (0.247)	-0.699 *** (0.248)
<b>Secondary Years</b>	0.884 *** (0.275)	0.252 (0.217)	0.250 (0.175)	0.235 (0.171)	0.021 (0.176)	0.047 (0.180)
<b>Price of Investment</b>	-0.007 * (0.003)	-0.013 *** (0.003)	-0.017 *** (0.003)	-0.015 *** (0.003)	-0.025 *** (0.005)	-0.022 *** (0.005)
<b>Gini (it)</b>	0.081 *** (0.029)	-0.029 (0.018)	-0.050 *** (0.014)			
<b>Avg. Gini (i)</b>				-0.073 *** (0.014)	-0.033 * (0.018)	-0.044 *** (0.020)
<b>Within-Gini (it)</b>				0.087 ** (0.034)	0.063 * (0.035)	0.067 * (0.035)
<b>Expropriation Risk</b>					0.770 *** (0.130)	0.817 *** (0.138)
<b>Polity in 1970</b>					--	-0.052 ** (0.02)
N	448	448	448	448	401	381
R-squared	0.55	0.04	0.08	0.12	0.22	0.24
Estimator:	FE	RE	Pooled	Pooled	Pooled	Pooled
% of error variance that is country specific:	82%	32%				

