Inequality and fractionalization

Casey, Gregory P. and Owen, Ann L.

Hamilton College

August 2010

Inequality and Fractionalization

Gregory P. Casey
Cornerstone Research
gregory.casey@fulbrightmail.org

Ann L. Owen*
Hamilton College
aowen@hamilton.edu

March 2012

Abstract: We present evidence that ethnic fragmentation explains variations in per capita income, institutions, and schooling better than income inequality. To do so, we identify instruments for ethnic fractionalization and income inequality based on historical experience and geography. Using instrumental variables estimation, we find that ethnic fractionalization explains the level of income both when income inequality is included as a control in the estimation and when it is not. However, we find inconsistent evidence that income inequality affects development when ethnic fractionalization is properly treated as an endogenous variable. If anything, the evidence we have suggests that income inequality is positively related to economic development.

Key Words: Inequality; Ethnic Fractionalization; Growth

JEL Codes: O11, O15, O16

*Corresponding author: 198 College Hill Road, Clinton, NY 13323; phone: 315-859-4419, fax: 315-859-4477

1 Introduction

A large literature finds a role for income inequality in affecting economic development both in the presence of credit market imperfections and without. The channels through which income inequality have been posited to work are numerous. For example, in a seminal paper, Galor and Zeira (1993) show that inequality can affect human capital accumulation-- and, therefore, growth-- in the presence of credit market imperfections. Galor and Moav (2004) extend this idea by arguing that inequality can be harmful in earlier stages of development when large investments in physical capital are primary drivers of growth but that the effects observed in Galor and Zeira (1993) may dominate later on. Moreover, Persson and Tabellini (1994) show how inequality can affect physical capital accumulation via a demand for redistributive policies, Alesina and Perotti (1996) argue that inequality affects physical capital investment through its effect on political instability, and Banerjee and Newman (1993) demonstrate a role for income inequality in affecting occupational choice and the extent of entrepreneurship. Others have linked inequality to the development of low quality institutions as the political elite block institutional reform that would benefit the country as a whole but challenge their own dominance (e.g. Acemoglu, Johnson and Robinson, 2001, 2005; Engerman and Sokoloff, 1997, 2000).[1]

At the same time, others have focused on the negative impact of a related but different aspect of societal division—ethnic fractionalization. Easterly and Levine (1997) show the negative consequences of ethnic diversity while paying particular attention to African development and argue that fractionalization interferes with the provision of growth promoting public goods. Others have confirmed the consequences of ethnic fractionalization (e.g. Alesina et. al. 2003), but Alesina and La Ferrara (2005) argue that ethnic diversity only has negative consequences in democracies where the lack of ability to coordinate across different ethnic groups may have more severe consequences. Importantly for our work, Alesina *et al* (2003) and Alesina and La Ferrara (2005) also argue that ethnic diversity is

[1] See Galor (2009a, 2009b) for brief but comprehensive treatments of the literature on inequality and growth.

endogenous and careful examinations of the role of ethnic diversity in affecting economic outcomes must take that into account. Ashraf and Galor (forthcoming) argue that genetic diversity can be beneficial at intermediate levels by spurring faster innovation but that the negative effects discussed here take over at higher levels, creating a U-shaped pattern.

In addition, Engerman and Sokoloff (1997, 2000) argue that ethnic diversity may have also played a role in the development of institutions by allowing the elites to readily identify a group that could be excluded from privileges such as landholding or suffrage. Thus, ethnic fractionalization may have negative impacts on development independent of the level of economic inequality. Indeed, their work highlights two potential roles of ethnicity in political development: it can be a tool for identification or a potential ideological fault line. Thus, *a priori*, it is unclear whether ethnic differences or income inequality are both playing independent roles in long run development.

In spite of a strong theoretical foundation for the effects of inequality on development, robust empirical evidence has been difficult to find. Clarke (1995) and Deininger and Squire (1998) find evidence that inequal ity is harmful for growth. Forbes (2000) finds that inequality has a positive effect on growth while Banerjee and Duflo (2003) find an inverted U relationship between growth rates and changes in the income distribution. Furthermore, they argue that it is difficult to interpret any of these correlations causally because of difficult identification problems. In an effort to address the identification issue, Easterly (2001a, 2007) uses characteristics of the land that might support plantation-based economies as instruments for inequality. He concludes that inequality is associated with lower levels of income, schooling and quality of institutions. Importantly, Easterly (2007) uses a variety of control variables, including ethnic fractionalization (treated as an exogenous variable) to demonstrate that income inequality is a robust determinant of the level of income per capita.

In contrast to the empirical literature relating income inequality to development, there is less debate about the negative consequences of ethnic fractionalization, at least in certain circumstances. The vast majority of this literature, however, treats fractionalization as exogenous. While treating ethnic fractionalization as exogenous may be an appropriate specification in classic growth regressions spanning 30 or so years, it is less acceptable over the longer time spans implicit in income level regressions (Alesina *et al*, 2003). Specifically, most migration occurs into those countries with higher levels of economic and institutional development, implying that OLS coefficients will underestimate the negative impacts of ethnic diversity (Mayda, 2005; Freeman, 2006). Moreover, ethnic diversity may hinder the development of institutions and provisions of public goods while countries are in the early stages of development, but have a smaller (or non-existent) negative impact once the institutions of democracy and rule of law have been implemented (e.g. Easterly, 2001b). Finally, a somewhat distinct literature demonstrates how national boundaries may be endogenous, which could create reverse causality between fractionalization and income per capita (e.g. Alesina and Spolaore, 1997; Alesina, Spolaore, and Wacziarg, 2005).

Recently, however, the treatment of ethnic fractionalization as an exogenous variable has begun to change. Michalopoulos (forthcoming) uses geographic variation to identify causes of fractionalization but does not relate fractionalization to income per capita. Similarly, Ahlerup and Olsson (forthcoming) appeal to an evolutionary framework, showing that ethnic fractionalization is positively related to an indicator of the duration of human settlement. Ahlerup (2010) directly addresses the endogeneity issue by treating fractionalization as endogenous when estimating its effects on income per capita and uses this framework to compare the effects of fractionalization and property rights institutions when both are treated as endogenous. His work does not, however, investigate the role of income inequality. In related work, Ashraf and Galor (forthcoming) use migratory distance from the location of human origins to predict genetic diversity and demonstrates an inverted U relationship between genetic diversity and income per capita.

We build on recent work by Putterman and Weil (2010) and Ahlerup and Olsson (forthcoming) on the historical determinants of inequality and ethnic fractionalization that identifies suitable instruments for examining the relative roles of ethnic fractionalization and income inequality. We find that ethnic fragmentation is an important determinant of per capita income, school enrollment and institutional quality. Most importantly, we demonstrate that, when income inequality and ethnic fractionalization are simultaneously added as endogenous variables in such regressions, ethnic diversity has a negative and significant effect while income inequality enters with a positive and usually statistically insignificant effect. These results are robust to numerous controls and different estimation strategies. The results also clearly indicate that fractionalization must be treated as an endogenous variable.

Our results are important because, although income inequality and ethnic fractionalization may be correlated empirically, the channels through which ethnic fractionalization affect economic development may be different than those through which income inequality affects economic growth. There are policy implications to this finding as well: the best policy to remedy the deleterious effects of ethnic fractionalization may be very different from one aimed at alleviating the effects of income inequality.

To the extent that ethnic fractionalization is a cause of income inequality or perhaps tells us something about the nature of the inequality, our results suggest that there are certain types of inequality that are worse for economic development than others. In other words, our results are consistent with the idea that inequality that is perpetuated by ethnic divisions may be particularly bad for economic growth. In that sense, to the extent that ethnic fractionalization affects growth via political or institutional channels rather than via factor accumulation, our findings complement those who argue that political inequality may lead to instability or lack of cohesion which lowers growth (e.g., Alesina and Perotti, 1996; Rodrik, 1999; Easterly, Ritzen and Woolcock, 2006). Also, our findings are consistent with the arguments in the strand of the inequality and growth literature that links inequality to the development of low quality institutions if exploiting ethnic divisions is a way for the elite to maintain their economic and political power when faced with growing domestic agitation for equal rights or when balancing inconsistencies inherent in arguing for freedom from colonizing powers while promoting the continuance of slavery (e.g. Engerman and Sokoloff, 1997, 2000).

2 Empirical Results

Our goal is to show that ethnic fractionalization and associated levels of inequality better explain poor growth performance than income inequality in general. To provide convincing empirical results, we will base our specifications on earlier empirical work demonstrating the link between inequality and growth (Easterly, 2001a, 2007). As mentioned in the introduction, our contention is that the forces that led to inequality also led to ethnic fractionalization. In particular, earlier literature attempts to identify geographic instruments that lead to inequality by providing incentives to develop plantations, which bred both inequality and ethnic divisions. Our departure from previous literature, which leads us to different conclusions, is that we will treat both fractionalization and inequality as endogenous variables.

2.1 Treating Fractionalization as an Endogenous Variable

We start by showing that previous results about the effect of inequality are not robust to treating ethnic fractionalization as endogenous. As mentioned above, Easterly (2007) uses land quality as an instrument for income inequality, treating ethnic fractionalization as exogenous. However, we show that, in addition to predicting income inequality, land quality also predicts ethnic fractionalization. The instrument used by Easterly (2007) is the likelihood that a country would export sugar or wheat. Specifically, the variable, *LWHEATSUGAR*, is defined as log (1+area of land suitable for growing wheat/1+area of land suitable for growing sugar). This data originally comes from the United Nations' Food and Agriculture Organization (FAO 2000). Easterly (2007) demonstrates that *LWHEATSUGAR* is significantly correlated with two measures of inequality: the percent of income controlled by the top 20% (*INCSHARE*) and the Gini coefficient (*GINI*), which are taken from the WIDER (2000) database. Both are averaged over the period of 1960-1998 in order reduce measurement error and are adjusted to account for biases introduced by different survey measurement techniques. Summary statistics, data descriptions and sources are provided for all the variables used in the paper in Appendix Table 1.

Columns 1 and 2 of Table 1 replicate the Easterly (2007) results for income inequality and Column 3 of Table 1 shows that we obtain similar results when we use *LWHEATSUGAR* to predict ethnic fractionalization.[2] The measure of fractionalization, originally from Alesina *et al* (2003), is the likelihood that two randomly selected individuals will be from different ethno-linguistic groups. The results in Table 1 suggest that a one standard deviation increase in LWHEATSUGAR decreases fractionalization, income share and the Gini coefficient by .37, .44, and .41 standard deviations respectively. The R-squared values for each simple regression in Table 1 are similar and, given the theoretical reasons to link land quality to both income inequality and ethnic fractionalization, this suggests that land quality may be just as good an instrument for ethnic fractionalization as it is for income inequality.

We show this point by presenting the estimation results of the effects of fractionalization and inequality when land quality is used as an instrument for each in separate regressions in Panel A of Table 2. Columns 1 and 4 present results when the Gini coefficient and the income share of the top 20 percent are used as measures of inequality and the dependent variable is the log of per capita income in 2002.[3] These results replicate those in Easterly (2007). However, in column 7 we report results when we substitute fractionalization for inequality and find that fractionalization also has the expected negative effect on income per capita.

As mentioned earlier, two ways in which inequality and fractionalization have been hypothesized to affect growth is via the accumulation of human capital or via the development of institutions. To see if there is evidence for these channels in our data we also report similar regressions in Table 2, using a measure of human capital accumulation and institution quality. So our results can be comparable to the previous literature we again use the same variables employed by Easterly (2007), *GOVERNANCE*, an aggregate measure of institutional quality from Kauffman, Kraay and Zoibo (2002) and secondary school enrollment rates averaged from 1998-2002 (SCHOOL). These results appear in

[2] We initially use the same sample as Easterly (2007) to show that our results are not due to a change in sample. Later, we are able to expand the sample.
[3] We use the log of income per capita in 2002 so that our results can be directly compared to those in Easterly (2007).

the remaining columns of Table 2 and show that in all cases, both inequality and fractionalization are significantly and negatively associated with institutional quality and schooling enrollments.

In Panel B of Table 2, we repeat the estimations results after adding the share of arable land in the tropics (*tropical*) as a second instrument in order test the over identification restrictions. Again, the explanatory variable is significant at the 1% level in all nine regressions and, more importantly, fractionalization passes the overidentification test in all three regressions, implying that the tests fails to reject the null hypothesis that *LWHEATSUGAR* and *tropical* affect the dependent variables through channels other than ethnic division. Importantly, diagnostic tests indicate that both inequality and fractionalization are endogenous. The test of endogeneity, which is in the last row of Table 2, is the difference between the Sargan-Hansen statistics when treating the potentially endogenous regressor as endogenous and exogenous. The null hypothesis is that the variable can be treated as exogenous and, in all specifications, this null hypothesis is soundly rejected for both inequality and fractionalization.

So, we have found that ethnic fractionalization and income inequality present equally plausible explanations for why the historical experience of countries as "plantation economies" is associated with slower growth than that experienced by their counterparts. Easterly (2007) demonstrates that his results for income inequality are robust to number of control variables, including commodity exporting indicator variable, the share of arable land in the tropics, legal heritage and continent dummies. We do not replicate his results in order to conserve space but do show similar estimations when ethnic fractionalization is treated as the endogenous variable. These results are in Table 3. In 11 of the 12 specifications, ethnic fractionalization is significant at the 5% level or better. When using income per capita as the dependent variable and adding *tropical* as a control, *FRAC* is significant at the 10% level (p-value of .06). Once again, diagnostic tests indicate that fractionalization should be treated as endogenous.

2.2 The "Horserace"

Because both fractionalization and income inequality appear to be important determinants of long-run development when investigated separately, we now enter them simultaneously into instrumental variables regressions. This "horserace" technique allows us to determine which variable exerts a greater effect and is similar to the approach taken by Acemoglu and Johnson (2005), Rodrik Subramanian and Trebbi (2004), and Easterly and Levine (2003) who attempt to determine the relative importance of competing factors that could influence growth. As we noted earlier, the two variables are correlated and the reason for similar results when using each one separately may be that one is proxying for the other. Of course, to employ a valid estimation technique, it is beneficial to have one instrument that predicts income inequality, but not ethnic fractionalization and one instrument that predicts fractionalization but not inequality.

Fortunately, recent literature on the historical determinants of current inequality and fractionalization provides us with such instruments. Putterman and Weil (2010) show that the standard deviation of state history (heterogeneity in the early development of the current population's ancestors) is related to current day income inequality.[4] We use this as an instrument for income inequality (in addition to the mean of adjusted state history as a control). To instrument for ethnic fractionalization, we rely on geographic features of the country, specifically, the natural logs of distance to a coast or a river and the absolute value of latitude. Ahlerup and Olsson (forthcoming) also use these geographic features as determinants of ethnic fractionalization. They argue that latitude is related to climatic variability, habitat diversity, and pathogen loads, which affect a local population's reach and migration, with greater variability encouraging the population to spread out more geographically and more pathogen's limiting migration in and out of a region. In addition, they suggest that populations that are closer to the coast or other waterways are less isolated and will be less fractionalized.[5] We use both distance to coast or river and latitude as instruments for fractionalization.[6]

[4] See Putterman and Weil (2010) for a detailed description of the construction of this measure. Underlying data is from Putterman, Louis. State Antiquity Index (available online) and Putterman, Louis and David Weil. World Migration Matrix, 1500 – 2000 (available online).

[5] Ahlerup and Olsson (2011) also explain linguistic fractionalization with a variable that captures the duration of uninterrupted human settlement, *origtime*. We use a different measure of fractionalization that incorporates both linguistic and racial diversity and do not find that *origtime* is a significant predictor of this measure of

As we explain below, although so far we have employed Two Stage Least Squares (2SLS) so that our results are comparable to previous literature, we will perform subsequent instrumental variables estimation using Limited Information Maximum Likelihood (LIML). Although this is a single equation technique, we can write the structural model as follows:

$$FRAC_i = \beta_{11}COASTRIVER_i + \beta_{12}LATITUDE_i + \beta_{13}SDSTATEHIST_i + \beta_{14}STATEHIST_i + \beta_{15}\boldsymbol{W}_i + \varepsilon_i$$

$$INEQ_i = \beta_{21}SDSSTATEHIST_i + \beta_{22}STATEHIST_i + \beta_{23}COASTRIVER_i + \beta_{24}LATITUDE_i + \beta_{25}\boldsymbol{X}_i + \omega_i$$

$$y_i = \gamma_1 FRAC_i + \gamma_2 INEQ_i + +\gamma_3 STATEHIST_i + \boldsymbol{\delta Z}_i + \mu_i$$

where *FRAC* is ethnic fractionalization, *INEQ* is either the income share of the richest quintile or the Gini coefficient, and y is the log of real income per capita in 2000,[7] secondary school enrollment rates (average of 1998 to 2002 values), or the Kaufman, Kraay, and Ziobo (2002) measure of good governance. As mentioned above, the instruments, COASTRIVER , LATITUDE, SDSTATEHIST are the natural log of distance to the coast or a river, the natural log of the absolute value of latitude, and the standard deviation of state history calculated by Putterman and Weil (2010). Consistent with our reasoning above, we expect $\beta_{13,}$ $\beta_{23,}$ and β_{24} to be equal to zero. Control variables in the estimation include regional dummy variables and dummy variables for legal heritage. As noted earlier in our discussion of Table 3, Easterly also controls for share of arable land in a tropical location and a commodity exporting dummy. We do not report these as part of our main specification because these variables do not enter the estimations significantly and their inclusion does not alter our qualitative conclusions.

---

fractionalization once latitude and distance from coasts or rivers are also included. Including *origtime* as an instrument does not change our conclusions, however.

[6] Our use of absolute value of latitude is similar to Easterly's use of the share of tropical land. We prefer to use latitude in our main specifications because it varies over a wide range of countries and is used in Ahlerup and Olsson (2011) to predict fractionalization.

[7] Switching to the log of GDP in 2000 rather than Easterly's measure for 2002 allows us to greatly expand the sample.

The specification above is over-identified. In this case, Limited Information Maximum Likelihood (LIML) provides coefficients that are less biased than Two Stage Least Squares (2SLS) if instruments are weak, although those estimates are less precise. Because of this, conclusions based on LIML results are more robust. Earlier, we presented results from 2SLS to provide the most direct comparison to previous literature. However, going forward, we use LIML for overidentified specifications.[8]

Although LIML is a single equation technique, we can still verify the expected relationships between inequality, fractionalization and the instruments described above. Results from the reduced form estimates of inequality and fractionalization are in Table 4. The first three columns examine the determinants of fractionalization, the middle three columns examine the Gini coefficients, and the last three columns present results for the income share of the richest quintile. These results are in line with the expectations we outlined above. The standard deviation of state history is a positive and statistically significant predictor of both measures of inequality, but not consistently of ethnic fractionalization. Conversely, the natural log of the distance to a coast or river and of the absolute value of latitude are significant predictors of ethnic fractionalization, but not consistent predictors of income inequality.

We report the results of our LIML estimations in Tables 5 and 6. In Table 5, the measure of income inequality used is the Gini coefficient. In Table 6, the income share of the top 20% is used. Prior to examining the main results, we first examine further specification tests that assure us of the validity of our estimation strategy. Specifically, in the bottom rows of Tables 5 and 6, we report several tests of underidentification and weak identification that support our efforts to separately identify the effects of inequality and fractionalization. The AP F-statistic is the Angrist-Pischke (2009) first stage F-statistic, which measures the strength of the instruments separately after controlling for the predicted values of the other endogenous variable. Low F-statistics indicate that we cannot reject the null

[8] As we explain later, we are able to reject the hypothesis of weak instruments in the LIML estimations for most of our specifications based on Stock and Yogo (2005) critical values. In addition, we find that overidentified 2SLS results are qualitatively similar to LIML results. Nonetheless, we report LIML estimates for overidentified models because of their superiority. (Angrist and Pischke (2009, p. 210).

hypothesis that the instruments do not separately identify the endogenous variables—i.e., the specification is underidentified. In all cases, we can comfortably reject this null hypothesis. A stronger specification test, however, would also test for weak identification. It is well known that weak instruments result in biased coefficients and very weak instruments can result in a bias that is as large as the OLS bias (Stock and Wright, 2000). However, using the Stock and Yogo (2005) critical values for a test of weak instruments based on LIML size, we find that in all but two specifications in Table 5, at the 5 percent significance level, the maximum size of the Wald test is less than .15, indicating that our estimation does not suffer from weak instruments.[9] To provide a more intuitive interpretation of the strength of our instruments, we also report the Shea's Partial R^2 for each of the endogenous variables in the first stage regressions. The adjusted Shea's partial R^2 takes into account the correlation between the instruments. Although there is much variation in each of these regressions that is unexplained, they do suggest that each endogenous variable can be separately identified with these instruments. Finally, Table 5 reports the results of a test of overidentifying restrictions, OIR, providing another measure of the validity of our instruments.[10] The results of all of these tests are particularly important in our context because we have two endogenous variables and the instruments may be correlated. However, in sum, these diagnostic results suggest that our instruments are valid for identifying separate effects of inequality and fractionalization.

We turn now to interpreting the main results of our instrumental variables estimation in Table 5. The first three columns examine income per capita, while the remaining six columns examine determinants of income per capita that may be influenced by inequality or fractionalization. Focusing first on the results for income per capita, we see that fractionalization has a strong negative effect on income per capita, while inequality actually has a positive effect that is statistically significant at the five percent level. These results show a meaningful role for fractionalization in determining income. For example, the results in column 1 predict that a one standard deviation increase in fractionalization (.25)

---

[9] The exceptions are in columns 2 and 8. However, even in these cases, the maximum size is estimated to be less than .20.
[10] We report p-values from the Hansen J-test for overidentifying restrictions.

would generate a 1.9 decrease in the log of GDP per capita. To determine if our results may be affected by omitted variables, we add legal heritage and continent dummies, the same control variables used by Easterly (2007), in columns 2 and 3. The magnitude of the coefficient for fractionalization decreases, but remains negative and significant at the one percent level. In contrast, the results for the Gini coefficient in the first three columns of Table 5 suggest that income inequality actually has a positive effect on income per capita.

The remaining columns of Table 5 examine the determinants of two variables through which income inequality or fractionalization may affect income per capita. In columns four through six, we report results of the estimation of the Kaufman, Kraay and Ziobo (2002) measure of good governance. As mentioned above, both income inequality and ethnic fractionalization have been theoretically linked to the development of weak institutions and poor governance. Our results find support for a negative link between fractionalization and poor governance, but we do not find evidence of the link between inequality and governance. This result is potentially important as it calls into question theories that link income inequality and weak institutions (e.g. Engerman and Sokoloff, 1997, 2000; Acemoglu, Johnson and Robinson, 2005). Our results suggest that it may be more appropriate to refine these theories to explicitly include elites that are identified by their membership in specific ethnic groups and not solely based on an income class. In columns seven through nine, we report estimation results for secondary school enrollments. Parallel to the results for income per capita, we find that income inequality is associated with higher enrollment rates and fractionalization is associated with significantly lower. The magnitude of the estimated effects of fractionalization are potentially devastating to economic development: the results in column 7 predict that a one standard deviation increase in fractionalization (.25) would decrease secondary school enrollment rates by 54 percentage points. As above, these results suggest that theories of access to education should specifically address membership in ethnic groups.

In Table 6, we replicate the Table 5 results using income share of the top quintile instead of the Gini coefficient as a measure of inequality. The conclusions from these alternative estimations are broadly consistent with those in Table 5; however, we are not as confident in our identification strategy when we use income share of the rich as the measure of inequality. Specifically, the Kleibergen-Paap statistic does not allow us to reject a maximum size of the Wald test as high as .20 in most specifications; therefore, we cannot conclude that this specification does not suffer from weak instruments. The difference in the two specifications, of course, is the measure of inequality and these results suggest that the standard deviation of state history is a better instrument for the Gini coefficient than for the share of income of the top quintile. Given that the standard deviation of state history applies to the entire population and that the Gini coefficient measures income inequality throughout the entire distribution while the income share of the top 20 percent only examines part of the distribution, this finding appears reasonable. Going forward, we will focus our analysis with the use of the Gini coefficient as the measure of income inequality. Nonetheless, the results we have obtained so far suggest that fractionalization is negatively associated with income, good governance, and secondary school enrollment rates. In contrast, income inequality is either positively related to these indicators of development or not related in a statistically significant way.

Why are our results different from those found by previous authors? The main difference in our approach is that we treat ethnic fractionalization as an endogenous variable in the estimation of income per capita. As we argued earlier, changes in income per capita can induce migration over the long-run which would affect the ethnic diversity of a country's population. If high incomes induced migration, treating ethnic diversity as exogenous causes the coefficient on ethnic diversity to be biased upwards, counteracting the hypothesized negative effect of ethnic fractionalization on income.

2.3 Fractionalization and Good Governance

To gain more insight into the ways in which inequality and fractionalization affect economic development, we provide more detailed results on the determination of specific aspects of good governance that comprise the overall Kaufman, Kraay and Ziobo (2002) measure. Results reported in Table 7 examine individual governance indicators: measures of "voice and accountability" (*VOICE*), "political stability" (*STABILITY*), "government efficiency" (*GOVEFFECT*), "regulation quality" (*REGQUAL*), "rule of law" (*RULELAW*) and "corruption" (*CORRUPT*). Fractionalization enters negatively and significantly in all 18 specifications, suggesting that ethnic diversity has a negative effect on all of the different aspects of "good governance" comprising the aggregate index. The estimation method used in Table 7 treats both fractionalization and inequality as endogenous in the determination of governance. This would be the case if good governance affects migration and therefore ethnic diversity and if governance affects income distribution. Both are plausible scenarios and, in fact, the tests for endogeneity (bottom rows of Table 9) indicate that the null hypothesis of exogenous fractionalization can be rejected at the 5% level in the majority of specifications. The null hypothesis for the Gini coefficient can be rejected in several, but fewer, specifications. Interestingly, income inequality is only statistically significant in one of these estimations and always estimated with a positive coefficient, suggesting that, after controlling for fractionalization, income inequality does not have a robust effect on governance.

3 Discussion

While this paper has focused on comparing the effects of inequality and ethnic fractionalization, we view our results as complementary to the growing literature on the effects of income inequality (e.g. Acemoglu, Johnson and Robinson, 2005; Engerman and Sokoloff, 1997, 2000; Easterly, 2007). Recent literature focuses on how inequality prevents the emergence of political systems that provide basic market-supporting institutions and public goods. The idea is that elites resist such institutions because they will allow the majority of the population to challenge their privileged position. Thus, the economic elites use their power to protect their own position at the expense of total economic growth.

A long literature on ethnic fractionalization suggests similar outcomes emerging through similar but distinct mechanisms. Specifically, different factions struggle to ensure that the allocation of government resources disproportionately benefit their own side. This struggle prevents countries from solving collective action problems and undercuts the validity and effectiveness of existing government institutions by spurring (in reality or perception) ethnic groups to use them in a partial manner. Thus, inequality prevents the emergence of high quality institutions because the people who control the development of their own political and economic institutions do not want to allow others to close the gap by improving their own position. Ethnic diversity, on the other hand, leads to a struggle between competing factions that can cause worse economic outcomes even if one group does not have complete control over the governmental resources. As in the inequality literature, the importance of ethnic diversity in long-run development works by interfering with the development of appropriate market-supporting institutions and the provision of public goods.

We provide evidence that ethnic fractionalization has a greater effect on economic development than income inequality. Although when the endogeneity of fractionalization is not considered, income inequality is associated with negative development outcomes, high income inequality is not a sufficient condition to lead to worse economic outcomes. In addition to showing that ethnic fractionalization affects the level of income, we also show that it affects investment in schooling. In our estimations, ethnic fractionalization has a strong negative effect on secondary school enrollment rates, while inequality has an insignificant and often positive impact. This result challenges the notion that elites will simply restrict access to public goods in order to maintain their privileged position. There are several reasons why elites may not block access to public education based simply on income differences. First, income differences could be insufficient motivation for a person to be willing to view the success of their group or (or himself) separately from the success of a country as a whole. Second, elites may simply lack the power to restrict public good provisions based on income. Finally, recent work in Unified Growth Theory argues that elites (or at least some portion of elites) benefit from having a more well-educated public (Galor and Moav, 2006; Galor, 2005; Golar, Moav and Vollrath, 2009). Instead, the results confirm the notion that ethnic differences prevent societies from overcoming the collective action problems inherent in providing public goods.

Another set of results suggests that ethnic fractionalization also effects income via the development of institutional quality. As above, we don't find that income inequality in general consistently affects institutions in a statistically significant way. The results for ethnic fractionalization, however, show many signs of persistent conflict in the development of effective institutions. These struggles manifest themselves in high levels of corruption and regulatory interference in the market and inefficient provision of services by the government bureaucracy and justice systems, all of which could results from ethnic groups attempting to use the government to extract rents from opposing segments of society. Similarly, lower scores on the "voice and accountability" measure could be the result of entrenched ethnic interests preventing real electoral competition. Finally, the relationship between ethnic fractionalization and "political stability and violence" is likely a response of the inability of diverse societies to mediate problems through established political channels.

4 Conclusion

We provide evidence that ethnic heterogeneity is better able to explain differences in income levels, school enrollment rates, and institutional quality than income inequality. Our results suggest that the nature of divisions in society may be particularly important in determining the effects of inequality on development.

We have provided evidence that ethnic fractionalization, rather than income inequality in general, is the major driving force behind differing paths to political and economic development. The results also suggest that ethnic differences have an important role to play in the literature on political transition (e.g. Engerman and Sokoloff, 2005; Acemoglu and Robinson, 2000, 2001, 2006). These results add to a long literature identifying the negative effects of ethnic fractionalization (e.g. Easterly and Levine, 1997; Alesina *et al*, 2003; Alesina and La Ferrara, 2005). Indeed, our results have implications for all regressions using the level of income as the dependent variable by implying that ethnic fractionalization must be treated as endogenous.

Works Cited

Acemoglu, Daron and Simon Johnson, 2005. “Unbundling Institutions,” *Journal of Political Economy* 113 (5): 949–995.

Acemoglu, Daron, Simon Johnson and James Robinson, 2001. “The colonial origins of comparative development,” *American Economic Review* 91 (5): 1369–1401.

Acemoglu, Daron, Simon Johnson, and James Robinson, 2005. “Institutions as the fundamental determinant of long term growth,” in Aghion, Philippe and Steven Durlauf (Eds.), *Handbook of Economic Growth*, Elsevier.

Acemoglu, Daron and James Robinson, 2000. “Why did Western Europe extend the franchise?: Democracy, inequality, and growth in historical perspective,” *Quarterly Journal of Economics* 115: 1167–200.

Acemoglu, Daron and James Robinson, 2001. “A Theory of Political Transitions,” *American Economic Review* 91 (4): 938–963.

Acemoglu, Daron and James Robinson, 2006. “De facto political power and institutional persistence,” *American Economic Review* 96 (2): 325–330.

Ahlerup, Pelle. 2010. “The effects of ethnic diversity: An instrumental variables approach,” Working Paper.

Ahlerup, Pelle and Ola Olsson, forthcoming. “The roots of ethnic diversity,” *Journal of Economic Growth*.

Alesina, Alberto and Eliana La Ferrara, 2005. “Ethnic Diversity and Economic Performance,” *Journal of Economic Literature* 43 (3): 762–800.

Alesina, Alberto and Roberto Perotti, 1996. “Income distribution, political instability, and investment,” *European Economic Review* 40 (6): 1203–1228.

Alesina, Alberto and Enrico Spolaore, 1997. “On the number and size of nations,” *Quarterly Journal of Economics* 112: 1027–1056.

Alesina, Alberto, Enrico Spolaore and Romain Wacziarg, 2005. “Trade, growth and the size of countries,” in Aghion, Philippe and Steven Durlauf (Eds.), *Handbook of Economic Growth*, Elsevier.

Alesina, Alberto, *et al*, 2003. “Fractionalization,” *Journal of Economic Growth* 8 (2): 155–194.

Angrist, Joshua and Jorn-Steffen Pichke, 2009. *Mostly Harmless Econometrics*, Princeton University Press.

Ashraf, Quamrul and Oded Galor, forthcoming, “Cultural diversity, geographical isolation, and the origin of the wealth of nations,” *American Economic Review*.

Banerjee, Abhijit and Esther Duflo, 2003. “Inequality and growth: What can the data say?” *Journal of Economic Growth* 8 (3): 267–299.

Banerjee, Abhijit and Andrew Newman, 1993. “Occupational choice and the process of development,” *Journal of Political Economy* 101 (2): 274–98.

Clarke, George, 1995. “More evidence on income distribution and growth,” *Journal of Development Economics* 47: 403–427.

Deininger, Klaus and Lyn Squire, 1998. “New ways of looking at old issues: Inequality and growth,” *Journal of Development Economics* 57: 259–287.

Easterly, William, 2001a. “The middle class consensus and economic development,” *Journal of Economic Growth* 6 (4): 317–336.

Easterly, William, 2001b. “Can institutions resolve ethnic conflict?” *Economic Development and Cultural Change* 49 (4): 687–706.

Easterly, William, 2007. “Inequality does cause underdevelopment: Insights from a new instrument,” *Journal of Development Economics* 84: 755–776.

Easterly, William and Ross Levine, 1997. “Africa's growth tragedy: policies and ethnic divisions,” *Quarterly Journal of Economics* 152 (4): 1203–1250.

Easterly, William and Ross Levine, 2003. “Tropics, germs, and crops: The role of endowments in economic development,” *Journal of Monetary Economics* 50 (1): 3–39.

Easterly, William, Jozef Ritzen and Michael Woolcock, 2006. “Social cohesion, institutions and growth,” *Economics and Politics* 18 (2): 103–120.

Engerman, Stanley and Kenneth Sokoloff, 1997. “Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the United States,” in Haber, Stephen (Ed.), *How Latin America Fell Behind*, Stanford University Press.

Engerman, Stanley, and Kenneth Sokoloff, 2000. “Institutions, factor endowments and paths of development in the New World,” *Journal of Economic Perspectives* 14 (3): 217-232.

Engerman, Stanley and Kenneth Sokoloff, 2005. “The evolution of suffrage institutions in the Americas,” *Journal of Economic History* 65 (4): 891–921.

Freeman, Richard, 2006. “People flows in Globalization,” *Journal of Economic Perspectives* 20: 145–170.

Food and Agriculture Organization (FAO), 2000. Global Agro-Ecological Zones. Available online.

Forbes, Kristin, 2000. “A Reassessment of the relationship between inequality and growth,” *American Economic Review* 90 (4): 869–887.

Galor, Oded, 2009a. “Inequality and Economic Development: An Overview,” Working Papers 2009-3, Brown University, Department of Economics.

Galor, Oded (Ed), 2009b. *Inequality and Economic Development: A Modern Perspective*, Edward Elgar Publishing.

Galor, Oded, 2005. “From Stagnation to Growth: Unified Growth Theory,” in Aghion, Philippe and Steven Durlauf (Eds.), *Handbook of Economic Growth*, Elsevier.

Galor, Oded and Omer Moav, 2004. “From physical to human capital accumulation: Inequality and the process of development,” *Review of Economic Studies* 71: 1001–1026.

Galor, Oded and Omer Moav, 2006. “Das Human-Kapital: A theory of the demise of the class structure,” *Review of Economic Studies*: 73 (1): 85–117.

Galor, Oded, Omer Moav and Dietrich Vollrath, 2009. “Inequality in landownership, the emergence of human-capital promoting institutions, and the Great Divergence,” *Review of Economic Studies* 76 (1): 143–179.

Galor, Oded and Joseph Zeira, 1993. “Income distribution and macroeconomics,” *Review of Economic Studies* 60: 35–52.

Harvard University, Center for International Development (CID). General Measures of Geography. Available online.

Kaufmann, Daniel, Art Kraay and Pablo Zoido-Lobaton, 2002. “Governance matters II - updated indicators for 2000-01,” Policy Research Working Paper Series 2772, The World Bank.

Mayda, Anna Maria, 2005. “International migration: A panel data analysis of economic and non-economic determinants,” IZA Discussion Papers 1590, Institute for the Study of Labor.

Michalopoulos, Stelios, forthcoming. “The origins of ethnolinguistic diversity: Theory and evidence," *American Economic Review*.

Persson, Torsten and Guido Tabellini, 1994. “Is inequality harmful for growth?” *American Economic Review* 84: 600–621.

Putterman, Louis. State Antiquity Index. Available online.

Putterman, Louis and David Weil, 2010. “Post-1500 population flows and the long-run determinants of economic growth and inequality,” *Quarterly Journal of Economics* 125 (4): 1627–1682.

Putterman, Louis and David Weil. World Migration Matrix, 1500 – 2000. Available online.

Rodrik, Dani, 1999. “Where did all the growth go? External shocks, social conflict, and growth collapses,” *Journal of Economic Growth:* 385–412.

Rodrik, Dani, Arvind Subramanian and Francesco Trebbi, 2004. “Institutions rule: The primacy of institutions over geography and integration in economic development,” *Journal of Economic Growth* 9: 131-165.

Stock, James and Jonathan Wright, 2000. “GMM with weak identification,” *Econometrica* 68 (5): 1055–1096.

Stock, James and MotohiroYogo, 2005, “Testing for weak instruments in linear IV regressions,” NBER Technical Working Paper 284.

World Bank. World Development Indicators, Various years.

World Institute for Development Economics Research (WIDER), 2000, World Income Inequality Database, Version 1.0.

Table 1: Land Endowments, Inequality and Fractionalization

	(1)	(2)	(3)
	GINI	INCSHARE	FRAC
LWHEATSUGAR	-18.33***	-19.13***	-0.441***
	(3.279)	(2.992)	(0.0961)
Constant	44.55***	49.28***	0.478***
	(0.923)	(0.798)	(0.0244)
Observations	118	114	118
R-squared	0.169	0.216	0.131

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Base Results with Inequality and Fractionalization Entering Separately

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	LNGDPPC02	GOVERNANCE	SCHOOL	LNGDPPC02	GOVERNANCE	SCHOOL	LNGDPPC02	GOVERNANCE	SCHOOL
Panel A: LWHEATSUGAR as Only Instrument									
GINI	-0.121***	-0.0914***	-4.891***						
	(0.0269)	(0.0200)	(0.960)						
INCSHARE				-0.127***	-0.0975***	-4.795***			
				(0.0292)	(0.0200)	(0.876)			
FRAC							-4.791***	-3.798***	-184.5***
							(1.050)	(0.819)	(33.97)
Constant	13.03***	3.910***	278.3***	13.89***	4.658***	296.8***	9.935***	1.657***	149.7***
	(1.132)	(0.847)	(39.44)	(1.388)	(0.951)	(40.60)	(0.483)	(0.359)	(14.52)
Observations	97	118	113	96	114	110	97	118	113
OIR									
FS F-Stat	27.42	31.23	28.80	31.24	40.90	37.74	20.49	21.05	25.48
FS P-Value	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
Endog	0.0001***	0.0002***	0.0000***	0.0001***	0.0001***	0.0000***	0.0017***	0.0007***	0.0001***
Panel B: LWHEATSUGAR and Tropical as Instruments									
GINI	-0.123***	-0.0961***	-4.933***						
	(0.0281)	(0.0208)	(0.981)						
INCSHARE				-0.128***	-0.0982***	-4.695***			
				(0.0297)	(0.0202)	(0.848)			
FRAC							-4.156***	-3.446***	-179.5***
							(0.856)	(0.703)	(33.22)
Constant	13.12***	4.117***	279.8***	13.94***	4.687***	291.7***	9.639***	1.501***	146.8***
	(1.187)	(0.880)	(40.35)	(1.411)	(0.961)	(39.28)	(0.393)	(0.315)	(13.93)
Observations	95	116	111	95	113	109	95	116	111
OIR	0.6308	0.5885	0.2188	0.3071	0.3022	0.0835*	0.3215	0.3320	0.8884
FS F-Stat	12.98	15.01	13.92	15.44	20.45	19.08	13.51	13.02	14.39
FS P-Value	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***	0.0000***
Endog	0.0000***	0.0001***	0.0000***	0.0000***	0.0000***	0.0000***	0.0054***	0.0007***	0.0001***

Estimated via 2SLS. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *OIR* is the p-value for the Hansen's J test of the exclusion restriction. *FS F-Stat* and *FS P-Value* are the F-Statistic and P-value from the first stage regression. Endog is the p-value from the test of endogeneity. Under the null hypothesis, the variable can be treated as exogenous. The test is calculated as the difference between the Sargan-Hansen statistics when treating the variable of interest as endogenous or exogenous.

Table 3: Effects of Fractionalization with Controls

	LNGDPPC02				GOVERNANCE				SCHOOL			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
FRAC	-6.561*	-4.479***	-5.852***	-4.406**	-5.369**	-4.274***	-4.888***	-4.047***	-189.0***	-182.3***	-211.2***	-132.4***
	(3.424)	(1.198)	(1.734)	(1.777)	(2.711)	(1.099)	(1.416)	(1.376)	(70.49)	(44.12)	(63.49)	(36.38)
Tropical	0.744				0.542				2.928			
	(1.045)				(0.767)				(21.72)			
Commodity		-0.238				0.298				-1.510		
		(0.295)				(0.290)				(12.56)		
British Heritage			0.784				0.612				31.73	
			(0.790)				(0.625)				(27.09)	
French Heritage			0.507				0.206				19.55	
			(0.643)				(0.522)				(23.22)	
Soc. Heritage			0.0745				-0.332				18.75	
			(0.636)				(0.457)				(18.27)	
Middle East/Africa				-0.348				0.176				-27.78**
				(0.558)				(0.401)				(10.93)
South & East Asia				-0.579				-0.299				-26.11***
				(0.365)				(0.253)				(9.232)
Western Hemisphere				0.118				0.195				-4.077
				(0.457)				(0.324)				(9.677)
Constant	10.43***	9.859***	9.894***	9.967***	2.163**	1.792***	1.935***	1.725***	150.0***	149.1***	139.7***	141.7***
	(1.191)	(0.502)	(0.556)	(0.533)	(0.949)	(0.439)	(0.397)	(0.458)	(24.69)	(16.79)	(14.53)	(12.05)
Observations	95	97	96	97	116	118	114	118	111	113	110	113
FS F-Stat	3.40	14.95	9.41	6.64	3.93	13.74	9.76	8.51	6.62	16.06	9.97	9.92
FS P-Value	0.0683*	0.0002***	0.0028***	0.0115**	0.0499**	0.0003***	0.0023***	0.0043***	0.0114**	0.0001***	0.0021***	0.0021***
Endog	0.0173**	0.0040***	0.0012***	0.0226**	0.0114**	0.0009***	0.0001***	0.0135**	0.0077***	0.0004***	0.0025***	0.0074***

Estimated via 2SLS. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *FS F-Stat* and *FS P-Value* are the F-Statistic and P-value from the first stage regression. Endog is the p-value from the test of endogeneity. Under the null hypothesis, the variable can be treated as exogenous. The test is calculated as the difference between the Sargan-Hansen statistics when treating the variable of interest as endogenous or exogenous. *LWHEATSUGAR* is the only instrument.

Table 4: Reduced Form Estimates of Inequality and Fractionalization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	FRAC	FRAC	FRAC	GINI	GINI	GINI	INCSHARE	INCSHARE	INCSHARE
SD of State History	0.181	0.324	0.108	34.61***	38.72***	30.20***	27.43***	25.06**	22.55***
	(0.236)	(0.316)	(0.236)	(8.485)	(12.81)	(8.567)	(7.968)	(11.83)	(7.763)
Ln(abs(latitude))	-0.101***	-0.104***	-0.0882***	-2.254**	-1.094	-1.024	-2.304***	-0.592	-1.047
	(0.0190)	(0.0239)	(0.0197)	(0.896)	(1.028)	(1.009)	(0.834)	(0.958)	(0.927)
Ln(Dist to a Coast or River)	0.0566***	0.0517***	0.0548***	0.231	-0.377	0.247	0.457	-0.233	0.508
	(0.0160)	(0.0173)	(0.0161)	(0.564)	(0.569)	(0.563)	(0.493)	(0.482)	(0.484)
State History	-0.109	-0.0390	-0.103	-7.586**	-4.420	-8.460***	-4.741*	-2.927	-5.681**
	(0.0947)	(0.0937)	(0.0943)	(3.003)	(3.425)	(2.868)	(2.777)	(2.961)	(2.703)
Middle East/Africa		0.0399			7.844***			9.635***	
		(0.0627)			(2.204)			(2.110)	
East and South Asia		-0.0730			0.585			3.430	
		(0.0618)			(2.476)			(2.177)	
Western Hemisphere		-0.0429			2.522			6.090**	
		(0.0680)			(2.897)			(2.713)	
British Heritage			0.213***			4.203*			3.147
			(0.0749)			(2.261)			(2.039)
French Heritage			0.156**			5.427***			5.164***
			(0.0683)			(1.839)			(1.751)
Socialist Heritage			0.142**			-0.311			-1.807
			(0.0672)			(2.024)			(1.878)
Constant	0.476***	0.477***	0.293**	48.07***	42.86***	41.86***	51.08***	43.99***	45.32***
	(0.129)	(0.130)	(0.130)	(5.080)	(5.363)	(5.491)	(4.567)	(4.934)	(4.959)
Observations	116	116	111	116	116	111	113	113	110
R-squared	0.362	0.381	0.413	0.336	0.426	0.387	0.288	0.426	0.364

Robust Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: The "Horserace" with the Gini Coefficient

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	LNGDPPC	LNGDPPC	LNGDPPC	GOVERNANCE	GOVERNANCE	GOVERNANCE	SCHOOL	SCHOOL	SCHOOL
GINI	0.121**	0.117**	0.144**	0.0334	0.0406	0.0294	3.305*	2.784*	3.982*
	(0.0548)	(0.0569)	(0.0690)	(0.0287)	(0.0370)	(0.0336)	(1.826)	(1.578)	(2.045)
FRAC	-7.421***	-4.328***	-6.824***	-3.935***	-3.097***	-3.827***	-216.5***	-102.8***	-190.8***
	(1.727)	(1.478)	(1.749)	(0.926)	(0.936)	(0.873)	(55.74)	(33.86)	(52.74)
Adjusted State History	1.145	1.692**	1.405	0.121	0.397	0.0605	29.43	37.96*	42.66
	(0.970)	(0.745)	(1.174)	(0.483)	(0.446)	(0.560)	(28.00)	(20.44)	(33.11)
Middle East/Africa		-1.575***			-0.476			-55.50***	
		(0.583)			(0.348)			(14.90)	
East and South Asia		-1.459***			-0.537**			-38.38***	
		(0.429)			(0.255)			(11.45)	
Western Hemisphere		-1.019			-0.402			-35.57**	
		(0.671)			(0.413)			(17.74)	
British Heritage			-0.518			-0.0931			-7.458
			(0.938)			(0.480)			(27.36)
French Heritage			-0.998			-0.446			-21.67
			(0.881)			(0.472)			(25.68)
Socialist Heritage			-0.682			-0.694*			6.982
			(0.628)			(0.361)			(18.10)
Constant	6.072***	5.582***	5.378*	0.250	-0.217	0.749	10.50	11.70	-26.06
	(2.147)	(2.110)	(2.887)	(1.096)	(1.322)	(1.396)	(69.21)	(59.57)	(84.92)
Observations	116	116	111	117	117	112	112	112	108
OIR	0.7212	0.7642	0.6613	0.8039	0.5138	0.8385	0.3713	0.4701	0.5179
AP F-Stat (GINI)	11.66	5.35	7.50	11.78	5.78	7.68	9.53	4.94	6.50
AP P-Value (GINI)	0.0000***	0.0061***	0.0009***	0.0000***	0.0041***	0.0008***	0.0002***	0.0089***	0.0022***
AP F-Stat (FRAC)	9.98	9.75	11.21	10.15	10.39	11.61	8.93	10.01	10.52
AP P-Value (FRAC)	0.0001***	0.0001***	0.0000***	0.0001***	0.0001***	0.0000***	0.0003***	0.0001***	0.0001***
Kleibergen-Paap F-Stat	5.26++	3.58+	4.94++	5.38++	3.90++	5.02++	4.50++	3.31+	4.18++
Shea's Partial R^2 (GINI)	0.1387	0.1046	0.1151	0.1403	0.1131	0.1178	0.1183	0.1013	0.1000
Adjusted Shea's Partial R^2 (GINI)	0.1156	0.0553	0.0640	0.1175	0.0647	0.0674	0.0938	0.0499	0.1898
Shea's Partial R^2 (FRAC)	0.1720	0.1848	0.2007	0.1731	0.1913	0.2022	0.1612	0.1919	0.0465
Adjusted Shea's Partial R^2 (FRAC)	0.1498	0.1399	0.1546	0.1511	0.1472	0.1566	0.1379	0.1458	0.1417
Endog (GINI)	0.0047***	0.0071***	0.0104**	0.0385**	0.0759*	0.0796*	0.0160**	0.0505*	0.0201**
Endog (FRAC)	0.0032***	0.0552*	0.0994*	0.6781	0.0215**	0.6152	0.1422	0.0597*	0.1911

Estimated via LIML. Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0. The instruments are the log of mean distance to a river or coast, the log of the absolute value of latitude and the standard deviation of Adjusted State History. AP F-Stats are the Angrist-Pishcke (2009) first stage F-Stats. OIR is the p-value of the test of the overidentification restriction. Endog is the p-value from the test of endogeneity. Under the null hypothesis, the variable can be treated as exogenous. The Kleibergen-Paap F-Stat is a joint test for weak instruments, which can be compared to the Stock and Yogo (2005) critical values. +++,++, and + indicate that, at the 5% significance level, the maximal size of the Wald test is less than .10, .15, and .20, respectively.

Table 6: The "Horserace" with the Income Share of the Richest Quintile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	LNGDPPC	LNGDPPC	LNGDPPC	GOVERNANCE	GOVERNANCE	GOVERNANCE	SCHOOL	SCHOOL	SCHOOL
INCSHARE	0.160**	0.179*	0.202*	0.0400	0.0577	0.0418	4.585	4.614	5.707*
	(0.0813)	(0.103)	(0.112)	(0.0382)	(0.0603)	(0.0500)	(2.975)	(3.156)	(3.459)
FRAC	-8.137***	-4.030**	-7.674***	-4.057***	-2.871***	-3.964***	-242.1***	-93.84**	-214.2***
	(2.241)	(1.576)	(2.185)	(1.085)	(0.914)	(1.008)	(79.47)	(39.39)	(67.30)
Adjusted State History	0.877	1.701**	1.261	0.0405	0.409	0.0503	23.97	40.76	40.35
	(1.053)	(0.862)	(1.322)	(0.484)	(0.474)	(0.581)	(31.37)	(25.55)	(38.10)
Middle East/Africa		-2.459**			-0.814			-80.22**	
		(1.020)			(0.566)			(31.87)	
East and South Asia		-2.032***			-0.774**			-53.68***	
		(0.613)			(0.322)			(18.23)	
Western Hemisphere		-1.854			-0.709			-59.15*	
		(1.159)			(0.654)			(35.51)	
British Heritage			-0.382			-0.0841			-4.963
			(1.044)			(0.494)			(30.64)
French Heritage			-1.143			-0.488			-25.97
			(1.082)			(0.517)			(31.60)
Socialist Heritage			-0.183			-0.557			20.88
			(0.791)			(0.404)			(23.44)
Constant	4.087	2.594	2.326	-0.114	-1.079	0.0946	-51.61	-76.29	-115.6
	(3.339)	(4.184)	(4.978)	(1.553)	(2.423)	(2.216)	(119.6)	(132.1)	(152.4)
Observations	113	113	110	114	114	111	110	110	107
OIR	0.6235	0.8281	0.7304	0.9576	0.5840	0.8227	0.4070	0.4861	0.6325
AP F-Stat (INC)	8.75	2.76	4.79	8.83	3.07	4.98	7.07	2.51	4.06
AP P-Value (INC)	0.0003***	0.0678*	0.0103**	0.0003***	0.0503*	0.0086***	0.0013***	0.0862*	0.0202**
AP F-Stat (FRAC)	7.67	10.35	8.75	7.77	11.12	9.21	6.56	11.23	8.15
AP P-Value (FRAC)	0.0008***	0.0001***	0.0003***	0.0007***	0.0000***	0.0002***	0.0021***	0.0000***	0.0005***
Kleibergen-Paap F-Stat	3.69+	1.87	3.08	3.79+	2.10	3.19	3.08	1.69	2.56
Shea's Partial R^2 (INC)	0.0933	0.0555	0.0712	0.0946	0.0628	0.0747	0.0769	0.0519	0.0605
Adjusted Shea's Partial R^2 (INC)	0.0683	0.0020	0.0170	0.0699	0.0103	0.0213	0.0507	-0.0033	0.0041
Shea's Partial R^2 (FRAC)	0.1354	0.1990	0.1629	0.1368	0.2070	0.1674	0.1199	0.2086	0.1553
Adjusted Shea's Partial R^2 (FRAC)	0.1116	0.1537	0.1141	0.1132	0.1625	0.1194	0.0950	0.1625	0.1046
Endog (INC)	0.0021***	0.0089***	0.0063***	0.0319**	0.1042	0.0645*	0.0110**	0.0697*	0.0127**
Endog (FRAC)	0.0896*	0.1098	0.3655	0.0071***	0.4816	0.5068	0.9745	0.2252	0.5025

Estimated via LIML. Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0. The instruments are the log of mean distance to a river or coast, the log of the absolute value of latitude and the standard deviation of Adjusted State History. AP F-Stats are the Angrist-Pishcke (2009) first stage F-Stats. OIR is the p-value of the test of the overidentification restriction. Endog is the p-value from the test of endogeneity. Under the null hypothesis, the variable can be treated as exogenous. The Kleibergen-Paap F-Stat is a joint test for weak instruments, which can be compared to the Stock and Yogo (2005) critical values. +++,++, and + indicate that, at the 5% significance level, the maximal size of the Wald test is less than .10, .15, and .20, respectively.

Table 7A: Breaking Down Institutions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	VOICE	VOICE	VOICE	STABILITY	STABILITY	STABILITY	GOVEFFECT	GOVEFFECT	GOVEFFECT
GINI	0.0729*	0.0580	0.0694	0.0328	0.0349	0.0387	0.0217	0.0494	0.0114
	(0.0399)	(0.0493)	(0.0447)	(0.0359)	(0.0435)	(0.0438)	(0.0343)	(0.0455)	(0.0392)
FRAC	-5.327***	-3.680***	-4.990***	-5.228***	-4.225***	-4.941***	-4.219***	-3.583***	-4.272***
	(1.327)	(1.287)	(1.233)	(1.164)	(1.127)	(1.109)	(1.099)	(1.154)	(1.028)
Adjusted State History	0.168	0.290	0.145	-0.437	-0.272	-0.339	0.449	0.909*	0.295
	(0.692)	(0.548)	(0.771)	(0.637)	(0.614)	(0.741)	(0.525)	(0.538)	(0.617)
Middle East/Africa		-0.763*			-0.558			-0.535	
		(0.454)			(0.456)			(0.419)	
East and South Asia		-0.728**			-0.466			-0.576*	
		(0.327)			(0.338)			(0.304)	
Western Hemisphere		-0.317			-0.385			-0.674	
		(0.514)			(0.499)			(0.516)	
British Heritage			-0.127			-0.112			0.0190
			(0.639)			(0.598)			(0.571)
French Heritage			-0.578			-0.414			-0.546
			(0.625)			(0.579)			(0.561)
Socialist Heritage			-0.787*			-0.213			-1.010**
			(0.477)			(0.441)			(0.440)
Constant	-0.848	-0.548	-0.420	0.963	0.700	0.778	0.720	-0.534	1.686
	(1.509)	(1.706)	(1.837)	(1.415)	(1.635)	(1.832)	(1.291)	(1.633)	(1.618)
Observations	117	117	112	117	117	112	117	117	112
OIR	0.7389	0.3507	0.9968	0.9125	0.6826	0.8950	0.7653	0.5905	0.7120
AP F-Stat (GINI)	11.78	5.78	7.68	11.78	5.78	7.68	11.78	5.78	7.68
AP P-Value (GINI)	0.0000***	0.0041***	0.0008***	0.0000***	0.0041***	0.0008***	0.0000***	0.0041***	0.0008***
AP F-Stat (FRAC)	10.15	10.39	11.61	10.15	10.39	11.61	10.15	10.39	11.61
AP P-Value (FRAC)	0.0001***	0.0001***	0.0000***	0.0001***	0.0001***	0.0000***	0.0001***	0.0001***	0.0000***
Kleibergen-Paap F-Stat	5.38++	3.90++	5.02++	5.38++	3.90++	5.02++	5.38++	3.90++	5.02++
Shea's Partial R^2 (GINI)	0.1403	0.1131	0.1178	0.1403	0.1131	0.1178	0.1403	0.1131	0.1178
Adjusted Shea's Partial R^2 (GINI)	0.1175	0.0647	0.0674	0.1175	0.0647	0.0674	0.1175	0.0647	0.0674
Shea's Partial R^2 (FRAC)	0.1731	0.1913	0.2022	0.1731	0.1913	0.2022	0.1731	0.1913	0.2022
Adjusted Shea's Partial R^2 (FRAC)	0.1511	0.1472	0.1566	0.1511	0.1472	0.1566	0.1511	0.1472	0.1566
Endog (GINI)	0.0099***	0.0759*	0.0207**	0.0864*	0.1720	0.1083	0.0954*	0.0585*	0.2030
Endog (FRAC)	0.0055***	0.0276**	0.2439	0.0197**	0.0311**	0.6730	0.0006***	0.0165**	0.9688

Table 7B: Breaking Down Institutions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	REGQUAL	REGQUAL	REGQUAL	RULELAW	RULELAW	RULELAW	CORRUPT	CORRUPT	CORRUPT
GINI	0.0543	0.0627	0.0457	0.0231	0.0518	0.0177	0.0224	0.0435	0.0202
	(0.0368)	(0.0494)	(0.0415)	(0.0328)	(0.0436)	(0.0388)	(0.0331)	(0.0459)	(0.0449)
FRAC	-4.755***	-3.463***	-4.506***	-4.276***	-3.703***	-4.294***	-4.253***	-3.733***	-4.550***
	(1.243)	(1.287)	(1.133)	(1.058)	(1.126)	(1.027)	(1.021)	(1.111)	(1.065)
Adjusted State History	0.284	0.626	0.197	0.435	0.949*	0.314	0.303	0.871	0.137
	(0.622)	(0.563)	(0.697)	(0.550)	(0.552)	(0.642)	(0.561)	(0.551)	(0.708)
Middle East/Africa		-0.751*			-0.482			-0.309	
		(0.436)			(0.408)			(0.430)	
East and South Asia		-0.754**			-0.620*			-0.719**	
		(0.308)			(0.321)			(0.327)	
Western Hemisphere		-0.576			-0.663			-0.503	
		(0.527)			(0.504)			(0.530)	
British Heritage			-0.113			-0.162			-0.214
			(0.580)			(0.550)			(0.624)
French Heritage			-0.518			-0.663			-0.641
			(0.564)			(0.528)			(0.610)
Socialist Heritage			-0.810*			-1.115***			-1.302***
			(0.443)			(0.412)			(0.463)
Constant	-0.307	-0.891	0.395	0.660	-0.645	1.507	0.713	-0.333	1.611
	(1.400)	(1.745)	(1.721)	(1.247)	(1.556)	(1.605)	(1.279)	(1.652)	(1.859)
Observations	117	117	112	117	117	112	117	117	112
OIR	0.6524	0.3425	0.9105	0.8484	0.6603	0.6197	0.9832	0.7551	0.3563
AP F-Stat (GINI)	11.78	5.78	7.68	11.78	5.78	7.68	11.78	5.78	7.68
AP P-Value (GINI)	0.0000***	0.0041***	0.0008***	0.0000***	0.0041***	0.0008***	0.0000***	0.0041***	0.0008***
AP F-Stat (FRAC)	10.15	10.39	11.61	10.15	10.39	11.61	10.15	10.39	11.61
AP P-Value (FRAC)	0.0001***	0.0001***	0.0000***	0.0001***	0.0001***	0.0000***	0.0001***	0.0001***	0.0000***
Kleibergen-Paap F-Stat	5.38++	3.90++	5.02++	5.38++	3.90++	5.02++	5.38++	3.90++	5.02++
Shea's Partial R^2 (GINI)	0.1403	0.1131	0.1178	0.1403	0.1131	0.1178	0.1403	0.1131	0.1178
Adjusted Shea's Partial R^2 (GINI)	0.1175	0.0647	0.0674	0.1175	0.0647	0.0674	0.1175	0.0647	0.0674
Shea's Partial R^2 (FRAC)	0.1731	0.1913	0.2022	0.1731	0.1913	0.2022	0.1731	0.1913	0.2022
Adjusted Shea's Partial R^2 (FRAC)	0.1511	0.1472	0.1566	0.1511	0.1472	0.1566	0.1511	0.1472	0.1566
Endog (GINI)	0.0171**	0.0561*	0.0405**	0.0999*	0.0451**	0.1983	0.1721	0.1035	0.3978
Endog (FRAC)	0.0000***	0.0283**	0.3178	0.0007***	0.0122**	0.8943	0.0005***	0.0097***	0.9796

Estimated via LIML. Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0. The instruments are the log of mean distance to a river or coast, the log of the absolute value of latitude and the standard deviation of Adjusted State History. AP F-Stats are the Angrist-Pishcke (2009) first stage F-Stats. OIR is the p-value of the test of the overidentification restriction. Endog is the p-value from the test of endogeneity. Under the null hypothesis, the variable can be treated as exogenous. The Kleibergen-Paap F-Stat is a joint test for weak instruments, which can be compared to the Stock and Yogo (2005) critical values. +++,++, and + indicate that, at the 5% significance level, the maximal size of the Wald test is less than .10, .15, and .20, respectively.

Appendix Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	Description	Source
LNGDPPC02	107	7.924	1.004	5.802	9.625	ln(GDP per capita in 2002)	Easterly (2007)
LNGDPPC	182	8.527	1.302	5.562	11.143	ln(GDP per capita in 2000)	WDI
SCHOOL	120	72.073	34.763	5.672	162.579	Average Secondary Schooling Enrollment Rates 1998-2002	Easterly (2007)
GOVERNANCE	128	0.085	0.784	-1.515	1.632	Aggregate Kaufmann, Kauffman, Kraay and Zoibo Governance Measure, 2002	Easterly (2007)
FRAC	127	0.427	0.246	0.002	0.930	Probability that two randomly selected individuals will be from different entho-linguistic group	Easterly (2007)
GINI	135	42.046	9.003	23.970	67.458	Gini Coefficient. Averaged 1960-1998.	Easterly (2007)
INC	129	46.640	8.687	17.573	71.211	Income Share of the Richest Quintile. Averaged 1960-1998	Easterly (2007)
SD of State History	161	0.097	0.089	0.000	0.346	SD of within-country migration-adjusted State History	Putterman and Weil (2010)
Ln(abs(latitude))	160	2.974	1.000	-0.862	4.314	Ln(abs(latitude)	Harvard CID
Ln(Dist to a Coast or River)	160	5.046	1.296	2.073	7.777	ln (mean distance to nearest coastline or sea-navigable river (km))	Harvard CID
LWHEATSUGAR	118	0.105	0.205	-0.393	0.578	log[(1+share of arable land suitable for wheat) / (1+share of arable land suitable for sugarcane)].	Easterly (2007)
Tropical	121	0.310	0.403	0	1	share of the country's cultivated land area in tropical climate zones	Easterly (2007)
Commodity	130	0.215	0.413	0	1	commodity exporting dummy	Easterly (2007)
State History	161	0.451	0.263	0	1	migration-adjusted State History	Putterman and Weil (2010)
Middle East/Africa	128	0.297	0.459	0	1	Continent Dummy. omitted category is Europe and Central Asia is the omitted cateogry	Easterly (2007)
East Asia	128	0.188	0.392	0	1	see above.	Easterly (2007)
Western Hemisphere	128	0.211	0.410	0	1	see above.	Easterly (2007)
British Heritage	122	0.303	0.462	0	1	Legal Heritage where German or Scandinavian legal origin are the omitted categories	Easterly (2007)
French Heritage	122	0.426	0.497	0	1	see above.	Easterly (2007)
Socialist Heritage	122	0.197	0.399	0	1	see above.	Easterly (2007)