Inequality, Redistribution, and Population

Filipe Campante
John F. Kennedy School of Government – Harvard University

Quoc-Anh Do
Department of Economics – Harvard University

October 2007

RWP07-046

Inequality, Redistribution, and Population[*]

Filipe R. Campante[†] and Quoc-Anh Do[‡]

June 2007

Abstract

We document a negative relationship between population size and inequality in the cross-country data. We propose an explanation built on the existence of a size effect in the political economy of redistribution, particularly in the presence of different channels of popular request for redistribution, e.g. "institutional" channels and "revolutions". Based on the assumption that the threat of revolution is directly related to the number of people that may attempt to revolt, the theory predicts that the stylized fact initially uncovered by the paper can be refined as follows: there is a negative relationship between population size, and its geographical concentration, and post-tax inequality in non-democracies. We subject these predictions to extensive empirical scrutiny in a cross-country context, and the data robustly confirm these patterns of inequality, population, and the interaction with democracy.

Keywords: Inequality, Redistribution, Population Size, Population Density, Population Concentration, Revolutions
JEL Classification: D31, D63, D74, J19

---

[*] This project was inspired by a discussion with Gregor Matvos, Demián Reidel and Jan Szilagyi. We owe special thanks to Philippe Aghion, Alberto Alesina and Andrei Shleifer for their guidance and their many helpful comments. We also acknowledge the useful suggestions given by Daron Acemoglu, James Alt, Robert Barro, Davin Chor, Jeffry Frieden, Torben Iversen, Michael Katz, Daniel Mejía, Emi Nakamura, Maria Petrova, Jón Steinsson, and seminar participants at the Economics and Government Departments at Harvard, the Political Economics Student Group Conference and the Association for Public Economic Theory 2006 Conference. The usual disclaimer on remaining errors certainly applies.

[†] John F. Kennedy School of Government, Harvard University; email: filipe_campante@ksg.harvard.edu

[‡] Department of Economics, Harvard University; email: do@fas.harvard.edu.

1 Introduction

This paper starts from the observation of a *negative* relationship between income inequality and population size in cross-country data: holding constant a number of usual explanatory factors for inequality, countries with larger population tend to be less unequal. To the best of our knowledge, this is a novel, and puzzling, observation. After all, why should there be any systematic relationship between these two variables, let alone a negative one? The intriguing nature of this empirical regularity suggests that explaining it can shed some light on the mechanisms and determinants of wealth and income inequality, and this is what the paper intends to pursue.

The existence of a size effect suggests the presence of some type of "increasing returns" in the process that generates the income distribution. While these increasing returns could be stemming from a variety of sources, we explore the possibility that they are linked to the political economy of redistribution. In fact, the political science literature has suggested that there is a significant size effect in political activities such as revolutions and insurgencies: Fearon and Laitin (2003) point out that insurgencies are more likely in countries with larger populations, because a large population "increases the number of potential recruits" for such activities (p. 81). They also provide empirical evidence consistent with that observation. From a slightly different theoretical angle, within the economics literature on revolutions, Grossman and Iyigun (1997) get to the same idea that larger populations are associated with more widespread subversive activities, for the return to these activities relative to production increases with population.

In order to investigate whether this is what is behind the empirical regularity, we build a very simple and stylized model that, similar to Acemoglu and Robinson (2000, 2005), analyzes the demand for redistribution when one of the channels for this demand is the threat of revolutions against a ruling elite. In that setup, we introduce a size effect and show that, if the threat of revolution is directly related to the number of people that may attempt to revolt – as distinct from the share of total population that this number corresponds to – then redistribution, which is partly driven by the desire to placate revolutionary threats, will be higher when population size is larger. The intuition is indeed very simple: the larger the population, the easier it will be to gather a mass of potential rebels that is enough to pose a serious threat of attempting a revolution. In particular, since poorer individuals have more incentive to demand redistribution, a larger population implies that this mass will typically be poorer. As a result, more redistribution will be required in order to contain this threat.

It is important to note that we do not need to assume that the *success* of a revolution attempt relies on the absolute number of supporters, as opposed to this number relative to that of opponents. Rather, what we need is that the *existence* of a serious revolution attempt be linked to the number of people that initiate it. This is because revolutions in our model are characterized by a "bandwagon effect" whereby, once under way, they can gain support well beyond the number of initial participants. (This is also a feature of much of the political science and sociology literature on the topic, e.g. Granovetter 1978, Kuran 1989, 1995.) By looking at specific historical evidence on a number of revolutionary episodes, we are able to provide additional support to the notion that absolute numbers are the key in setting off revolutionary attempts. Quantitative cross-country evidence also strengthens the case: the occurrence of revolutions and coups is positively related with population.

The upshot of our exercise is that this simple theoretical framework generates four basic additional predictions that can be tested so as to assess whether increasing returns in the politics of redistribution is what is behind the empirical link between inequality and population size. First, since the theory focuses on redistribution, it predicts that such link should not be present in the data on "gross" (pre-tax) inequality. Second, the theory implies that the effect should be less important in democratic countries: the revolution channel through which the size effect operates should be less salient in the demand for redistribution, as the latter can be voiced through democratic channels. Third, the model predicts a negative effect of population density, and of its concentration around the capital city, on inequality. This follows from the basic logic behind the theory, once it is assumed that it is harder to join a revolution attempt when one is farther away from the focal point of political turmoil. Put simply, more disperse populations will thus have a harder time putting together a revolutionary threat. (In that regard, the model is consistent with a lot of casual evidence on the concern with revolutionary strife that influences some countries' decisions to locate their political capitals distant from the main population centers, or dispersed in more than one location.) Finally, the model also predicts that the effect should typically decrease as size increases, since the level of redistribution required to prevent revolutions increases at a diminishing rate.

We provide extensive empirical evidence, using cross-country data, that supports all four predictions. In fact, they prove to be remarkably robust, as we control for the quality of the data and potential outliers, and take into account a number of additional control variables and the different types of inequality that are measured.[1] They are also robust to the use of different measures of inequality (e.g. Gini coefficient and

---

[1] On this topic, see the critique by Atkinson and Brandolini (2001).

the ratio of income held by top and bottom deciles) and of the geographical concentration of population (e.g. density and a new measure of population concentration around the capital city, which we have developed elsewhere, in Campante and Do 2007). In addition, our findings also hold when estimation is performed using dynamic panel techniques, which help to allay concerns over endogeneity arising from reverse causality or country-specific unobservables.[2]

In light of our theory, the initial observation of an empirical regularity linking population size and inequality should thus be qualified as follows: *There is a negative relationship between population size, and its geographical concentration, and post-tax inequality in nondemocratic countries.* Quite importantly, these additional predictions would be hard to reconcile with other possible explanations for the stylized fact under consideration, which strongly suggests that the political economy of redistribution is indeed at the heart of this empirical observation. In addition, this qualification helps us make sense of the instances that do not fit the coarser description of a negative relationship between size and inequality: examples that come to mind are large and unequal countries such as Brazil and South Africa (which have very low levels of concentration around the capital cities), and the comparison between the United States and Western European countries (which are by-and-large democratic).

Related Literature

Understanding the determinants of wealth and income inequality and redistributive policies is of course a first-order topic in economics that has received the attention of an extensive literature. Among the many proposed determinants and covariates are income and growth (as in the famous Kuznets hypothesis), ethnic fractionalization (see Alesina and Glaeser 2004), social mobility (Bénabou and Ok 2001; Alesina and La Ferrara 2005), political institutions and factor endowments (Engerman and Sokoloff 1997), etc. Population has not been prominent among those.

The study of the determinants and effects of the size of countries has gained momentum recently (see Alesina and Spolaore 2003and references therein). Within this literature, the relationship between inequality and size has been studied by Bolton and Roland (1997), but they take the former to be a long-run determinant of the latter, and in effect it is unclear from their argument which empirical relationship there should be in steady state, if any, between inequality and size. Our paper, on the other hand, follows the approach of using population as an explanatory variable (similar to Mulligan and Shleifer 2005 on regulation, for instance).

Some even more recent work has taken a somewhat skeptical position with respect to this literature on

---

[2] For an extensive coverage of these methods see Arellano and Honoré (2001).

size. Rose (2006) looks for evidence of scale effects on a host of economic variables of interest, including inequality measured by the Gini coefficient, at the cross-country level; he concludes that there is little evidence of any scale effect in addition to the well-recognized one on openness. His result is clearly consistent with our evidence, and in fact highlights one of the main messages of this paper: The effect of size on inequality depends on institutional features such as the level of democracy, and on which type of inequality one is dealing with.[3] Without recognizing these conditionants it is unsurprising, in light of our results, that no systematic relationship emerges.

Other papers in that literature, while not dealing directly with inequality, have focused on the role of government. Rodrik (1998), for instance, finds a positive relationship between openness and government size in the cross-country data, and attributes this relationship to a greater demand for social insurance, to be provided by government expenditure, in countries more subject to external shocks. Alesina and Wacziarg (1998) approach the question from a different angle, arguing that smaller countries have larger governments, as a share of GDP, because of economies of scale. They suggest that this is the reason behind that positive relationship. In either case, one might be tempted to extend the argument to redistribution, and predict that smaller countries would display more redistribution. Our evidence suggests that this conclusion would not be warranted; quite to the contrary, we observe that smaller countries actually redistribute *less*.[4]

This discussion also relates to the literature on "small states" in political science. This literature has the book by Katzenstein (1985) as its most recognizable starting point, where he argues that small states (in Europe) tend to have better social protection, because of their greater openness and exposure to volatility. This would also go in the opposite direction of what we find here, though not in an inconsistent manner. Our model predicts – and the empirics confirm – that the negative relationship between size and inequality should not be present in more advanced democracies, which are the focus of Katzenstein's analysis. Others have tried to empirically assess the relationship between smallness and inequality – as in Brautigam and Woolcock (2001) and references therein – with mixed results. However, these investigations are more limited in scope than what we propose here, since they focus on a restrictive definition of smallness, as opposed to looking at a continuous measure of size. In that regard, we show that our results are not driven by the very small countries this literature typically looks at.

---

[3] Indeed, Rose's (2006) data on inequality comes from the CIA World Factbook, which does not make clear whether the compiled Gini coefficients refer to net or gross income.

[4] Alesina and Wacziarg (1998) do provide some evidence that can be seen as consistent with ours: in their Table 8, they obtain a positive coefficient of population on public expenditure, significant at the 10% level, when the measure is inclusive of transfers.

The rest of the paper is organized as follows: Section 2 presents the basic stylized fact; Section 3 presents the evidence in support of our crucial assumption on revolution attempts; Section 4 develops the theoretical framework; Section 5 contains the empirical tests of this framework; and Section 6 discusses possible alternative explanations for the stylized fact. Section 7 concludes.

2 A Stylized Fact on Inequality and Population Size

This section presents a preliminary empirical investigation of the relationship between inequality and population size. In particular, we focus on the cross-country variation in size and investigate whether it is a relevant predictor of inequality, measured by the Gini coefficient. We utilize the most up-to-date and complete inequality dataset, the World Income Inequality Database (WIID) version 2.0 assembled by the World Institute of Development Economic Research (WIDER). This is a much revised and updated version of the WIID 1.0, which built on Deininger and Squire's (1996) dataset. Inequality datasets are usually criticized for their lack of consistency both across and within countries (Banerjee and Duflo 2003, Atkinson and Brandolini 2001). The new version of the WIID goes a long way in addressing much of that criticism by carefully considering the characteristics of the surveys leading to each observation, and classifying them under several categories. Notably, it is made clear whether each survey conveys information on income or expenditure, what form of income or expenditure is concerned, and whether the concepts and methodology of the survey are clear and reasonably correct.

We analyze only the observations that, in a quality scale of 1 (highest quality) to 4, attain an index of 1 (clear income concepts, reliable and verifiable surveys) or 2 (either the income concept or the survey is verifiable). This step assures the quality of the data under analysis, and is similar to the use of only "acceptable/reliable quality" data from earlier datasets (WIID 1, Deininger-Squire). Very importantly, unlike the latter, WIID 2.0 does not restrict the high-quality sample to have a single observation per country-year, which enables us to keep track of different kinds of sources, e.g. consumption, income, expenditure etc.[5] This is in line with the recommendation of Atkinson and Brandolini (2001) in their critique of the Deininger-Squire dataset. We then contract the dataset to represent only one observation for each country, year and type of data. As suggested by Deininger and Squire (1996) and reconfirmed by Atkinson and Brandolini (2001), we include a dummy variable for consumption-based inequality

---

[5] More specifically, we define three categories to encompass all the income definitions (labeled "incdefn" in the WIID data set) that are present in the data. "Consumption" includes incdefn = "Consumption", "Consumption/Expenditure" and "Expenditure"; "Gross" includes incdefn = "Earnings, Gross", "Income, Factor", "Income, Gross", "Income, Taxable", "Market Income", and "Monetary Income, Gross"; "Net" includes incdefn = "Earnings, Net", "Income, Disposable", and "Monetary Income, Disposable".

data in all regressions. We also acknowledge the fundamental difference between gross and net income inequalities, a point to which we will come back later, and use another dummy variable to indicate to which kind of inequality data each observation refers.[6]

We will use mostly the logarithm of population (referred to interchangeably as size) as our main explanatory variable, taken from the World Development Indicators (WDI). The logarithmic transformation helps alleviate the overwhelming importance of big countries in the dataset, as noticed in Yitzhaki (1996), and also in interpreting the results as average percentage changes in population, not average absolute changes. (Our results are even stronger when we use the level, rather than the log of population.) The final dataset includes 1395 observations covering 104 countries, ranging from 1960 to 2003, with good quality information on inequality categorized by types of measure. (Table A1 in the appendix describes the sample.) Because different countries typically end up with a different number of observations in the final dataset, there is the concern that results might be driven by those countries that are strongly represented in the sample. In order to alleviate that concern, we use weighted least squares throughout the analysis, giving equal weight to each country in the sample.

The simplest starting point is to look at Figure 1, which plots inequality against the log of population, where inequality is the simple country average of the sample over time, ignoring distinctions between types of data. There is already a hint of a negative relationship, but no clear pattern emerges. Table 1 establishes the basic stylized fact in the context of a regression analysis. Although the simple regression of inequality on size displays essentially no effect (Column (1)), as soon as we include the first set of controls a strongly significant *negative* relationship emerges. We successively add different control variables, starting with GDP per capita in PPP terms (from the WDI) and openness (measured as the share of total imports and exports in GDP, from the WDI) in Column (2); then ethnolinguistic fractionalization (from Alesina et al. (2003)), the effectiveness of democratic institutions (the "Polity 2" variable in the Polity IV dataset), and the logarithm of land area (from the WDI) in Column (3). The inclusion of land area checks the consistency of the result insofar as land area is a natural predictor of population size, whereas the other four variables address different channels of impact on inequality previously identified in the literature, as mentioned in the Introduction. The negative size effect on inequality remains strongly significant.

---

[6] Fixed effect panel regressions show that consumption based Gini coefficients are on average 2.2 points lower than the non-consumption based ones, which is nonnegligible, although less than the number 6.6 suggested by Deininger and Squire (1996) for expenditure-based data. A similar exercise yields a difference of 1.9 points between gross and net income inequality. Those numbers are statistically significant at the confidence level of 0.1%.

[FIGURE 1 HERE]

[TABLE 1 HERE]

Since each country typically has more than one observation in the sample, it is also important to consider specifications with error clustering at the country level. Column (4) shows that the result is robust to clustering as well. We finally add a few extra sensitivity checks in Columns (5)-(7), still keeping the clustered errors – the inclusion of dummies for legal origin, which could control for systematic differences between countries with different historical backgrounds (for instance, socialist countries typically display lower inequality), region dummies, to control for unobserved regional specificities that may affect inequality, and a dummy for China and India, which could be driving the result since they are outliers in terms of population, and relatively equal. The coefficient on size remains negative across all the specifications, and strongly significant with the exception of the specification with regional dummies. In words, larger countries are likely to be less unequal than smaller ones.

The coefficients are also economically significant. For instance, with a coefficient of -1.5, the difference in population size between countries in the 25th and 75th percentiles of the size distribution – which corresponds to comparing a country of 5 million inhabitants (e.g. Kyrgyzstan in the early 2000s) to one of 59 million (e.g. Egypt in the mid-1990s) – predicts a difference in the Gini coefficient of about one half of a standard deviation (controlling for differences due to the various types of inequality data).

The evidence of this robust negative relationship between population size and inequality calls for further theoretical investigation. First of all, it cannot be simply a mechanical artifact, since the Gini coefficient is conceptualized to be independent of scale, in the sense that it is invariant to the replication of the population (e.g. Bourguignon 1979). It also contradicts the political science literature that predicts the opposite relationship (e.g. Katzenstein 1985), which typically argues that smaller countries, being more exposed to external shocks, must provide better social insurance, including via redistribution, leading to lower inequality.

We will present a theory that accounts for the stylized fact we uncover, based on the presence of economies of scale in the political economy of redistribution. More specifically, this explanation builds on the interaction of different channels through which redistributive pressures are exerted, one of which being the possibility of "revolutions" – by which we mean the overthrowing of an incumbent ruler or elite. As will be seen, it takes as its basic premise – the one that gives rise to those economies of scale – the assumption that what is crucial for starting a revolution attempt is the absolute number of individuals involved, rather than the share of total population it represents. Therefore, we start by providing a discussion of the adequacy of this assumption.

3 Revolutions and Numbers

> "It is not admissible that fifty individuals in the Republic's capital be able to unsettle and threaten fifty million Brazilians."
>
> Juscelino Kubitschek, President of Brazil (1956-60)[7]

3.1 Historical Evidence

When talking about the Ukrainian revolution of 2004/05, *The Economist* states that "Kiev's key lesson [on revolutions] is that numbers are all-important: 5,000 or even 15,000 people can be violently dispersed; 50,000 are a different proposition." (March 18th 2006, p. 28) We can draw two important elements from this statement: first, numbers are crucial in revolutions; second, 50,000 is a rather small proportion of the 47 million Ukrainians in 2004. This aptly illustrates our main idea: *absolute* numbers are the main factor in putting together a revolution attempt. The larger the number of people involved in starting a revolutionary episode, the more likely it is to be important; however, this number need not represent a large proportion of total population. We take this idea from the political science literature on insurgencies, where Fearon and Laitin (2003) claim that larger populations are positively associated with insurgencies because they imply a larger source of rebels, and also that "the number of active rebels (...) is often in the hundreds or low thousands." (p. 81) In other words, a few thousand people may be a negligible proportion of the population of a given country, but a few thousand people in the streets of the capital could under some conditions start a chain reaction that eventually brings down the government. To illustrate that, let us consider a few historical examples in some more detail.

A first example is that of the Russian Revolution in 1917. According to the Tacitus Historical Atlas, the population of what was to become the Soviet Union was by then around 184 million people, and that of Russia proper was around 100 million.[8] However, by all accounts, the revolution was started by a tiny proportion of this large population. Trotsky himself notes that "the [February] revolution was carried out upon the initiative and by the strength of one city [Petrograd], constituting approximately about 1/75 of the population of the country." (1932, ch. 8) Moreover, even within Petrograd it was a far smaller number that actually produced the revolutionary spark, by starting a series of strikes that

---

[7] Quoted in Couto (2001). The translation is our own.

[8] The figure is 93.5 million people in 1926, but the data for the Soviet Union indicate a sharp drop in population between 1917 and 1926.

culminated in the fall of the czarist regime: Trotsky (1932, ch. 7) estimates that about 90,000 people, less than 0.01 percent of the country's population, took part on the first day of strikes.[9]

Once the movement got started, it rapidly gained momentum, with increasing numbers taking on to the streets of Petrograd, and eventually took down the government a mere four days after the first strike. After the revolution was under way, the number of forces willing to fight for the regime, which was in great disarray as a result of the Russian participation in World War I, vanished rapidly, to the point that "by night [the czarist garrison, numbering 150,000 soldiers] no longer existed." (Trotsky, 1932, ch. 7) This example thus clearly illustrates how revolutions are started by a small share of the population and then gain momentum quite rapidly, given favorable conditions.

Another example comes from the French Revolution. Although pinpointing a specific moment for the onset of this revolutionary episode is a much more difficult task, the one that is usually agreed on as such is the fall of the Bastille, in July 14, 1789. The *vainqueurs de la Bastille*, as designated by the French National Assembly in June 1790, were 954 people, but the size of the Parisian mob that stormed the fortress that stood as a symbol of the Ancien Regime should perhaps be estimated in five to six thousand people, as cited in contemporaneous newspaper accounts quoted by Spielvogel (1999, p. 416). Le Bon (1913, p. 173) goes as far as an estimated 12,000 people, but in any case the fact is that the numbers are hardly impressive as a share of the 27 million inhabitants that France is estimated to have had at the time. Once again, this event in which a tiny proportion of the population took part arguably put in motion a chain of events that – at a much more uneven pace than that of the Russian episode – ended up overthrowing the existing regime.

More recent episodes can also be brought to bear on that topic. One of those is that of East Germany in the late 1980s and early 1990s. Lohmann (1994) identifies the "Monday demonstrations" that took place in the city of Leipzig during the fall of 1989 as the event that started the political turmoil that culminated with the toppling of the East German regime and the ensuing German reunification, in October 1990. As Lohmann (1994, p. 69) points out, just over six thousand people took part in the first of these demonstrations, and some eighteen thousand took part in the second. For the sake of comparison, the total population of East Germany was by then around 16 million people, according to the figures of the Federal Statistical Office of Germany. Lohmann (1994) also documents how these figures swelled into the hundreds of thousands in Leipzig, and into the millions in the country as a whole, once again

---

[9] Our examples are meant to be confined to events that represent a real attempt at revolution, in the sense that they are linked to a revolutionary episode in an immediate way. That excludes, for instance, the founding of the Bolshevik Party, or similarly less-than-immediate factors.

illustrating how rebellions initiated by a relatively small number of people grow to the point where the regime in power is brought down.

Yet another modern revolutionary episode can provide an example coming from a smaller country: the Sandinista revolution in Nicaragua, in the late 1970s. In that case, the Library of Congress Country Study on Nicaragua (Merrill 1994) claims that by 1978 "hard-core Sandinista guerrillas numbered perhaps 2,000 to 3,000; untrained popular militias and foreign supporters added several thousand more to this total". The country's population by then was just under 3 million, according to INEC, the country's official statistics provider. The aforementioned country study goes on to state that soon the Sandinista front "no longer was fighting alone but rather was organizing and controlling a national insurrection of citizens eager to join the anti-Somoza movement", in a pattern reminiscent of the other examples discussed before. Another example from a small country is a more contemporaneous one: the events that culminated with the overthrowing of the government of Kyrgyzstan in March 2005. As an article in *The Economist* pointed out, "only a few thousand of Kyrgyzstan's 5m people took to the streets" (*The Economist,* March 31st, 2005), yet the regime was toppled in a few days' time.

A few points stand out as lessons to be taken from these examples, for our purposes. First, the evidence seems very supportive of the idea that the number of people that is sufficient to start a revolution attempt need not represent a significant share of the population. One does not need millions, or even hundreds of thousands, to start a revolution in countries with millions of inhabitants. Second, these examples, with the exception of the Russian one, involve similar numbers, of a few thousand, even though the population in the countries involved varies by a factor of ten. In other words, the population share that the initial rebels correspond to is significantly higher in smaller countries.

Another important point is that the claim here does not refer to the number of people that eventually join a revolution attempt, but only to the number of people that start one. Similarly, it is not about the success of a revolution attempt; we do not claim that the French or Russian Revolutions succeeded with only a few thousand people. The point is that revolutions can be started by a number of people that, while small when compared to the relevant population, is enough to start off a chain of events in which more and more people are drawn onto the revolutionary "bandwagon". This "bandwagon" or "domino" effect is indeed consistent with existing models of revolutions (Granovetter 1978; Kuran 1989, 1995; Lohmann 1994 and references therein), as well as with accounts of revolutionary episodes, as suggested by the few examples briefly presented above.

In sum, the historical evidence is quite supportive of the assumption that it is absolute numbers, and not numbers relative to total population, that matter in determining the emergence of a revolutionary attempt.

3.2 Empirical Evidence on Revolutions and Population

We can complement the lessons obtained from the historical record by looking at some quantitative empirical evidence. After all, if our hypothesis is correct, we would expect a positive impact of population on the occurrence of revolutionary episodes. Fearon and Laitin (2003) provide evidence in that direction, in the context of civil wars: they show that population is a strong predictor of the emergence of conflicts. We provide our own basic test for that using a measure of the average number of revolutions per year across countries in the period between 1960 and 1984, from the Barro-Lee data set. We then proceed to investigate whether there is any connection between this variable and population size. Note that, since we have a single average value for each country, we also average the variables used in the previous section over our sample.

Table 2 displays in general a positive correlation between population size and the number of revolutions in all specifications, indicating that countries with larger population experience more revolutionary episodes. The coefficient is significantly positive except for the last two specifications, where the inclusion of log land area and openness drops the value of the estimated coefficients of population size and renders them less precise. This can be attributed to the high correlation between these two variables and population size. Nevertheless, these results are reasonably consistent with our hypothesis, and together with the previous results in the literature we take them to be suggestive evidence.

[TABLE 2 HERE]

4 A Very Simple Model of Revolution and Redistribution

Let us now present a very simple framework to understand the interaction between population and redistribution. While far from a full-fledged theory of revolutions, it nevertheless enables us to explore the consequences of the presence of increasing returns in the political economy of redistribution.

4.1 Basic Framework

Consider a country with a population of size N. At the beginning, the country has a certain level of imperfect democracy that gives unequal rights to its citizens – for example, wealthier individuals may have more political power. The wealth of citizen i is denoted w_i, and total wealth is distributed according to a cdf $F(w)$. There is a government (or "elite") that has measure zero, and which redistributes and keeps the country running. We assume that this elite gets an exogenous payoff from being in power, and zero if overthrown.

Each citizen, given her wealth w, has a desired level of tax rate $\tau^d(w)$. We assume without directly modeling that there is a non-degenerate, monotonic distribution of those tax rates (such that $\tau^d(w_1) > \tau^d(w_2)$ if $w_1 < w_2$).[10] Given inequalities in political power, the political process in place implements a given tax rate, τ_v. In sum, this is the first channel that determines redistribution, which works through the "conventional" political process.

This process, however, can be disrupted by "revolution".[11] Dissenting citizens may choose to attempt a rebellion to overthrow the government. In case such an attempt happens, there is a fixed social cost K that everyone suffers, but every individual can choose whether to join the revolt, to fight against it, or to stay passive. If the revolution attempt succeeds, those who directly participated in it expect a private gain b_1, whereas those who fought against it expect a punishment of p_2. Similarly, if it fails, those who take part in the successful counterrevolution expect a private gain of b_2, and the participants in the failed attempt get a punishment of p_1. Those who stay aside are neither punished nor rewarded.[12]

After a successful revolution, society is assumed to reach perfect democracy, under which political power is equally divided among citizens. We do not mean that all revolutions will lead to democracy, which would obviously fly in the face of all historical evidence, but rather to have a simple way of capturing the idea that a primary motive for revolutions is redistribution, and that they often replace the existing regime with a more broadly-based one (Bueno de Mesquita et al. 2003, ch. 8).[13] The implemented tax rate will be then τ_m (the median voter's preferred rate), which we assume to be such that $\tau_m \geq \tau_v$ – meaning that the initial political process is biased towards the rich, as consistent with the evidence discussed in Bénabou (2000). Equality happens only when the country was already a perfect democracy in the first place. Finally, if the revolution attempt fails, the payoffs will still correspond to the tax rate of τ_v. The timing of the model is shown in Figure 2.

[FIGURE 2 HERE]

Let us solve the model using backward induction. Let the probability of success of a revolution once

---

[10] The reader could refer to Alesina and Rodrik (1994) for a detailed model of such assumption.

[11] Our model is similar to Acemoglu and Robinson (2000, 2005). They build a story of democratization spurred by the threat of revolution, while we build a story of redistribution taking democracy to be exogenous.

[12] It can be shown that the results still hold if the punishment and reward are assumed to depend on the number of agents joining or fighting against the revolution attempt, provided that the punishment is bounded.

[13] In fact, Roemer (1985) derives the commitment to progressive redistribution as the optimal behavior of a rational revolutionary leader.

it is attempted be denoted π. Let us assume that it depends on the number of people taking part in the revolution, n_f, and on the number of people resisting it, n_a: $\pi = \pi(n_f, n_a), \pi_1 > 0, \pi_2 < 0$. We also impose $\pi(0, \cdot) = 0$ (i.e. if nobody joins the attempt, it will surely fail), and $\pi(N, 0) \equiv \overline{\pi} > 0$ (if all the masses join, there is a positive probability of success, though we leave open the possibility that the elite still manages to stave off an attempt joined by all the masses).[14] The payoffs for each of the options are:

$$\pi(n_r, n_a)\left[(1-\tau_m)\, w_i + \tau_m \overline{w} + b_1\right] + (1 - \pi(n_r, n_a))\left[(1-\tau_v)\, w_i + \tau_v \overline{w} - p_1\right] - K, \tag{1}$$

$$\pi(n_r, n_a)\left[(1-\tau_m)\, w_i + \tau_m \overline{w} - p_2\right] + (1 - \pi(n_r, n_a))\left[(1-\tau_v)\, w_i + \tau_v \overline{w} + b_2\right] - K, \tag{2}$$

$$\pi(n_r, n_a)\left[(1-\tau_m)\, w_i + \tau_m \overline{w}\right] + (1 - \pi(n_r, n_a))\left[(1-\tau_v)\, w_i + \tau_v \overline{w}\right] - K. \tag{3}$$

It follows that an individual will join the revolution if (1)>(2) and (1)>(3):

$$\begin{aligned} \pi(n_r, n_a) b_1 - (1 - \pi(n_r, n_a))\, p_1 \quad &> \quad -\pi(n_r, n_a) p_2 + (1 - \pi(n_r, n_a))\, b_2 \implies \\ &\implies \quad \pi(n_r, n_a)\, (b_1 + p_1 + b_2 + p_2) > p_1 + b_2, \end{aligned}$$

$$\pi(n_r, n_a) b_1 - (1 - \pi(n_r, n_a))\, p_1 > 0 \implies \pi(n_r, n_a)\, (b_1 + p_1) > p_1.$$

Note that, although the decision to start a revolution will be shown to depend on the agent's wealth, the decision to join *once the revolution is already taking place* does not. It follows that, if the rewards from membership in a successful revolution or the punishment from membership in a failed counter-revolution are high enough, there exists an equilibrium in which everyone joins an ongoing uprising. In this case, revolutions display the bandwagon effect we alluded to in the previous section.[15] We focus on the case where everyone joining is an equilibrium, and we assume that if this is the case, this equilibrium is selected.[16] The conditions for this to be an equilibrium are: $\overline{\pi}\,(b_1 + p_1 + b_2 + p_2) > p_1 + b_2$, and $\overline{\pi}\,(b_1 + p_1) > p_1$. For the model to be interesting, we assume that we are in a situation where these conditions hold.

Assumption 1: $\overline{\pi}(b_1 + p_2) > (1 - \overline{\pi})(p_1 + b_2)$, and $\overline{\pi} b_1 > (1 - \overline{\pi}) p_1$.

---

[14] We could also consider the case in which $\overline{\pi}$ is a function of N. For example, we could have $\overline{\pi}'(N) < 0$, meaning that the elite's ability to fight a revolution attempt is increasing with population size, perhaps because of economies of scale in military technologies (see Alesina and Spolaore 2003). We will later address the implications of this alternative assumption.

[15] This is a stronger result than what we actually need for the model's main points. For example, one could imagine that the punishment for engaging in a failed counter-revolution (p_2) depends on the individual's wealth – though it is not clear that this should be the case, since this is the punishment beyond the post-revolution redistribution. In any case, this assumption would impose a participation cutoff in the wealth distribution. The crucial feature is that participation would be unrelated with gaining or losing from redistribution, so the bandwagon effect would still be present. On the other hand, this would lead to the relevant $\overline{\pi}$ being smaller, which, as will be seen, would reduce the range in the parameter space over which a revolution could plausibly be attempted.

[16] Note that there always is an equilibrium where no one joins the revolution, and mixed-strategy equilibria, where each individual is indifferent between joining or not, are also possible. There is thus a non-trivial problem of equilibrium selection that a full-fledged theory of revolutions would have to address. For our purposes, it suffices to establish the possibility of the bandwagon effect.

Note that if we have $\overline{\pi} = 1$, this boils down to having $b_1 > 0$ or $p_2 > 0$, that is to say: if membership in a successful revolution entails any benefit, or membership in a failed counter-revolution effort entails any punishment, everyone joining is indeed an equilibrium.[17]

Going back one stage, each individual has to decide whether it is in her interest to attempt a revolution. Individual i's payoff after redistribution at tax rate τ is $(1-\tau)w_i + \tau\bar{w}$, where $\bar{w}$ is the population average of wealth. Knowing that an attempted revolution will be successful, the individual will find it in her interest to attempt a revolution if and only if:

$$\overline{\pi}\left[(1-\tau_m)w_i + \tau_m\bar{w} + b_1\right] + (1-\overline{\pi})\left[(1-\tau_v)w_i + \tau_v\bar{w}\right] - K \geq (1-\tau_v)w_i + \tau_v\bar{w}. \tag{4}$$

After simplification, this condition is equivalent to:

$$w_i \leq \bar{w} - \frac{K/\overline{\pi} - b_1}{\tau_m - \tau_v}.$$

The right hand side defines the threshold under which the individual would choose to rebel. Not surprisingly, it is the poor who have an incentive to revolt, but the threshold is lower if the social cost of the revolution is higher, or if the initial system is close to democratic. More precisely, we have:

$$w^* \equiv \max\left(0, \bar{w} - \frac{K/\overline{\pi} - b_1}{\tau_m - \tau_v}\right). \tag{5}$$

The total number of people willing to attempt a revolution is thus given by:

$$n_r = F(w^*)N. \tag{6}$$

Here we impose the crucial assumption in our framework, as extensively discussed in Section 3:

Assumption 2: For a revolution to be attempted, a *critical mass* $\overline{n}_r$ of individuals willing to do so must be attained.

Consider the case $F(w^*)N \geq \bar{n}_r$ (equivalent to $w^* \geq F^{-1}\left(\frac{\bar{n}_r}{N}\right) \equiv \Omega$, with Ω increasing in $\bar{n}_r$ while decreasing in N). This is the condition for the existence of a revolutionary equilibrium. It is easy to see that this condition is more likely to be met when $N, \overline{\pi}, \tau_m$ are higher, or when K, τ_v are lower. The intuitions are clear: When population is larger, it is easier to motivate a number of rebels that is large enough for a successful attempt; by the same token, when the cost of revolution is lower, dissenting citizens find it easier to revolt. The comparative statistics on τ's involves the aspiration of democratic redistribution: when the aspired redistribution rate is higher, or the current redistribution rate is lower,

---

[17] The importance of the existence of private benefits for revolutionary success is emphasized in Bueno de Mesquita et al. (2003, p. 372).

dissenters are more willing to revolt. Finally, if the elite's ability to fight a revolution attempt once it is under way is lower, as captured by a higher $\overline{\pi}$, the willingness to revolt also increases.

The elite, which is assumed to be overthrown by a successful revolution, will try to redistribute at the start of the game in order to avoid the rebellion.[18] In other words, it will try to adjust the tax rate to a level that prevents a revolution to be attempted. While the elite always prefers to avoid revolutions, we can also assume that increasing τ_v also entails some cost. Therefore, the elite would choose the lowest "pacifying" rate, τ_v^P, such that no revolution attempt will arise. Since the condition for an attempt is $w^* \geq F^{-1}\left(\frac{\bar{n}_r}{N}\right) \equiv \Omega$, using (5) it is easy to see that the elite will set:

$$\tau_v^P = \tau_m - \frac{K/\overline{\pi} - b_1}{\bar{w} - \Omega}. \tag{7}$$

As Ω is decreasing in N, τ_v^P is increasing in N. This implies our main result:

Proposition 1 *If $K/\overline{\pi} > b_1$(i.e. the social cost of a revolution is sufficiently high), then in an imperfect democracy that starts with $\tau_m - \tau_v \geq \frac{K/\overline{\pi} - b_1}{\bar{w} - \Omega}$, redistribution increases with population size N. Consequently, the post-fiscal distribution in larger countries will be less unequal.*

Proof. *It suffices to compute the derivative of τ_v^P with respect to N:*

$$\frac{\partial \tau_v^P}{\partial N} = \frac{K/\overline{\pi} - b_1}{(\bar{w} - \Omega)^2} \frac{1}{F'(\Omega)} \frac{\overline{n}_r}{N^2} > 0,$$

given the assumption on $K/\overline{\pi} > b_1$. ■

The intuition here is once again transparent. For a given pre-tax distribution, a larger population will mean that the critical mass that is necessary to attempt a revolution can be reached at a lower level of wealth. As a result, more redistribution will be required in order to prevent a revolution from taking place. Note that ours is a model of *potential*, not actual, revolutions (as Acemoglu and Robinson 2000, 2005): revolutions do not occur on the equilibrium path. In other words, in our model it is *not* the case that larger countries are more equal because they experience more revolutions, but because their larger revolutionary potential requires more redistribution in order to be contained.[19]

A key testable prediction immediately comes out of Proposition 1: the size effect will only hold if democracy is sufficiently imperfect, where the imperfection is measured by the distance $\tau_m - \tau_v$. This

---

[18] Grossman (1995) discusses general conditions under which it is optimal for the elite to redistribute in order to prevent the masses from engaging in "extra-legal appropriation".

[19] If we consider the case where $\overline{\pi}$ is a function of N, which we have mentioned above, the result of Proposition 1 is maintained under some reasonable conditions. The details are left for the Appendix, but the intuition is as follows: If there are diseconomies of scale in the counter-revolution technology ($\overline{\pi}'(N) > 0$), the size effect on redistribution is reinforced. If there are economies of scale ($\overline{\pi}'(N) < 0$), on the other hand, we have a force on the opposite direction, but the result holds provided that the economies of scale are small enough or, less restrictively, that they decrease fast enough.

enables us to derive a first test of our proposed explanation: *The effect of population size should hold only for nondemocratic countries.* The mechanism that our theory proposes is turned off in democracies, where redistributive demands find room within the institutionalized political process. In light of that, our theory is indeed consistent with the aforementioned argument advanced by Katzenstein (1985), which refers primarily to democratic countries.

We can also derive a statement about the second derivative of redistribution with respect to population size:

Corollary 2 *If the pre-tax distribution $F(\cdot)$ is not too convex at Ω, then redistribution is concave in population size N: as population increases, redistribution increases at a diminishing rate.*

Proof. Computing the second derivative yields:

$$\begin{aligned}\frac{\partial^2 \tau_v^P}{\partial N^2} &= -2\frac{K/\overline{\pi}-b_1}{(\bar{w}-\Omega)^3}\frac{1}{[F'(\Omega)]^2}\frac{\overline{n}_r^2}{N^4}+\frac{K/\overline{\pi}-b_1}{(\bar{w}-\Omega)^2}\frac{1}{[F'(\Omega)]^3}F''(\Omega)\frac{\overline{n}_r^2}{N^4}-2\frac{K/\overline{\pi}-b_1}{(\bar{w}-\Omega)^2}\frac{1}{[F'(\Omega)]}\frac{\overline{n}_r}{N^3}= \\ &= \frac{K/\overline{\pi}-b_1}{(\bar{w}-\Omega)^2}\frac{1}{[F'(\Omega)]}\frac{\overline{n}_r}{N^3}\left[\frac{1}{F'(\Omega)}\frac{\overline{n}_r}{N}\left(F''(\Omega)-\frac{2}{(\bar{w}-\Omega)\,F'(\Omega)}\right)-2\right].\end{aligned}$$

It is easy to see that this expression is negative unless $F''(\Omega)$ is positive and too large. ■

Intuitively, the key is that as N increases the "pivotal" individual that needs to be coopted is at a lower point in the distribution, but this move occurs at a diminishing rate. Provided that the pre-tax distribution is not too convex, this implies that the rate at which the wealth of the pivotal individual decreases is also diminishing, and the required redistribution will then also increase at a diminishing rate. In other words, Assumption 2 implies a "fixed cost" in the revolution technology. This leads to a second testable prediction: *The effect of population size on inequality should be weakened as population size increases.*

4.2 The Role of Population Density and Concentration

4.2.1 An Extension

The crucial assumption in our theory is that the *number* of potential rebels is what matters for a revolution attempt to take place, not the *share* of the population they represent. However, this brings up an immediate counter-argument: if those rebels are scattered across the countryside, they should not be that menacing. To deal with that objection, we can extend our basic framework as follows: Assume that the population is unevenly distributed over the territory. For simplicity, assume that there is a "center" of population, which one can think of as being the country's capital city where political decisions are made. Finally, assume that the individual cost of joining the rebellion depends (linearly for simplicity) on the agent's distance from that center, δ_i, meaning that there is a cost of "getting together" at the capital city to plot a revolution.[20] Then we can rewrite (4) as:

$$\overline{\pi}\left[(1-\tau_m)w_i + \tau_m\bar{w} + b_1\right] + (1-\overline{\pi})\left[(1-\tau_v)w_i + \tau_v\bar{w}\right] - \delta_i - K \geq (1-\tau_v)w_i + \tau_v\bar{w}.$$

This can be simplified to:

$$w_i + \frac{\delta_i}{\overline{\pi}(\tau_m - \tau_v)} \leq \bar{w} - \frac{K/\overline{\pi} - b_1}{\tau_m - \tau_v}.$$

If we define $w^*(\delta) \equiv w^* - \frac{\delta_i}{\overline{\pi}(\tau_m - \tau_v)}$, we have the new revolution participation threshold, which obviously decreases with distance: the farther you are from the center, the poorer you must be so that it compensates to join the revolt. This can be seen in Figure 2, where $\delta^* = w^*\overline{\pi}(\tau_m - \tau_v)$. What matters now is the joint probability distribution $G(w, \delta)$, and the equivalent of (6) is now:

$$n_r = N \int_0^{\delta^*} \int_0^{w^*(\delta)} g(w,\delta)\,dw\,d\delta. \tag{8}$$

Following the same reasoning from the previous subsection, we now have the additional pacifying redistribution τ_v^P being defined implicitly by:

$$\overline{n}_r = N \int_0^{\overline{w}\overline{\pi}(\tau_m - \tau_v^P) - (K/\overline{\pi} - b_1)} \int_0^{\overline{w} - \frac{K/\overline{\pi} - b_1 + \delta/\overline{\pi}}{(\tau_m - \tau_v^P)}} g(w,\delta)\,dw\,d\delta.. \tag{9}$$

This once again shows that an increase in N will increase redistribution τ_v^P, by the same logic as before. However, a key variable here is the dispersion of $g(\cdot,\cdot)$ along the δ dimension: if we decrease this dispersion, it is clear from Figure 3 that we will be increasing the number of people that will find it optimal to attempt a rebellion. The consequence will be increased redistribution in equilibrium.

[FIGURE 3 HERE]

The crucial lesson of this extension of the theoretical framework is that the spatial distribution of the population matters: in short, potential rebels who are close to the capital city are more dangerous to an incumbent regime than those who are far away. One way to give empirical content to this prediction is to focus on *population density*: for a given population size, increasing density would correspond to a decreased dispersion.

The main advantage of this approach is that there are readily available empirical measures of population density. However, keeping constant population size, changes in population density only speak

---

[20] This "distance" could perhaps be interpreted in terms of some other dimension, e.g. ethnic fractionalization, along which individuals are unevenly distributed. This could provide some other mechanism through which fractionalization leads to redistribution, which is unlike the usual argument: instead of redistributing because people find it more appealing to redistribute to those who are like them, here redistribution would occur because more homogeneous individuals would find it easier to coordinate in a revolt, increasing the risk of a revolution.

to a narrow range of comparisons across land area measures. A more thorough inspection of equation (9) suggests that the prediction would refer more precisely to some concept of *population concentration* around the capital city. An adequate concept of concentration around a center is what we need to capture geographical changes in the distribution $G(w, d)$. We will discuss such a concept later, in the empirical implementation.

In sum, we are able to provide another set of testable predictions coming from this extended framework: population density and population concentration should have positive effects on redistribution in non-democratic countries, as they should be associated with an increased revolutionary threat.

4.2.2 Discussion

This crucial role of the geographical concentration of the population within the logic of our model invites a few considerations. First, this model actually formalizes one of the mechanisms highlighted by Alesina and Glaeser (2004) in explaining why the United States redistributes much less than Western Europe. They note that "America's vast geographic spread ensured that despite the dramatic success of many early labor groups in the United States, it was impossible to organize an effective nationwide movement that threatened the entire nation."(Alesina and Glaeser 2004, p. 107) They also proceed to observe how important rebellions could not gather enough momentum so as to topple the national government due to the distance between their epicenters in major population centers such as New York and the political capital in Washington, DC. These factors in turn greatly reduced the pressures for redistribution, as suggested by our theory.

The example of the United States indeed illustrates the consequences for redistribution of having a capital city that is distant from the main population centers. In terms of our model, the capital city can be thought of as the relevant "focal point"for a revolution; after all, as put by *The Economist*, "during a [revolutionary] stand-off, the capital city is crucial."(March 18th 2006, p. 28) Along those lines, choosing the location of the capital is actually a way of manipulating population concentration. The ruler can reduce concentration, and thereby alleviate the revolutionary pressures, by choosing an isolated location as the site of the capital.[21] It is not hard to come up with examples in which the model's logic is displayed rather transparently. In the 17th century, Louis XIV moved away from the Parisian masses into the tranquility of Versailles, a move that is widely thought to have been influenced by his having

---

[21] The relationship between population of "central"cities and political instability has been explored by Ades and Glaeser (1995). They focus on the main population center, rather than the capital city, since they are interested in investigating non-political factors for city growth as well. However, they do emphasize the role of proximity to the political center and political influence through revolutionary threats. They also consider the possibility of moving the capital city away from population centers to manipulate political incentives.

witnessed and suffered the rebellions against the Crown that became known as the *Frondes* (1648-53): "The Fronde made a lasting impression on Louis, creating a lifelong fear of rebellion. It also left him with a distaste for Paris, the largest city in his kingdom and the center of royal government."(MSN Encarta). Modern examples are also easy to come up with: countries such as Brazil (1960), Nigeria (1991), Kazakhstan (1997), and most recently Myanmar (Burma) (2005), have moved their capitals from important population centers to newly built cities in sparsely populated areas.[22] Many other countries that have not done it have fiddled with the idea.[23] In just about every case, a chief concern was to have the new capitals to be "quiet, orderly places where civil servants could get on with their jobs without distraction."(*The Economist*, Dec. 18th 1997) And when moving the capital city is not really a possibility, incumbent governments pay a high price for not reinforcing their base of support within the capital. The recent military coup in Thailand on Sep. 19th, 2006 is a striking example of how a government could be overthrown effortlessly if devoid of support from the population of the capital city, even when such government was largely popular in the countryside (*The Economist*, Sep. 22nd, 2006).

Looking closely at one of these examples helps illuminate the connection with our theory. For instance, Brazil had the capital moved in 1960 from Rio de Janeiro to Brasília – about 1,000km from the main population centers of Rio de Janeiro and São Paulo, and far from the coast, where most of the country's population was and still is. The debate over moving the capital is much older, though, and from the start the advantages of moving away from the crowds were acknowledged by those in favor of the idea: as early as 1810, while Brazil was still under Portuguese rule, an advisor to the king made the point that "the capital should be in a healthy, agreeable location free from the clamorous multitudes of people indiscriminately thrown together." (*The Economist*, Dec. 18th 1997) As Couto (2001) remarks, the president who finally decided to build the new capital, Juscelino Kubitschek, was also guided by a desire to escape from the atmosphere of political agitation in Rio, where the president was more exposed to political crises and student demonstrations. As his quote in the epigraph of Section 3 suggests, Kubitschek was instinctively aware of the logic that our model uncovers.

Of course, many other factors can play a role in the decision to move the site of government – tensions across different regions have certainly played a role in cases such as the United States and Nigeria, as have concerns about vulnerability to foreign attacks in many other cases. But it is clear nonetheless that protection against rebellion has typically been a prominent concern. In that regard, it is interesting to

---

[22] The Burmese move to the fortified "secret mountain compound"is an illuminating, if somewhat extreme example. (*International Herald Tribune*, Nov 11th 2005)

[23] See the article in *The Economist* (Dec. 18th, 1997) for a discussion of many examples.

note that two countries that are prominent examples of acute potential for redistributive strife, Brazil and South Africa, have chosen either to locate the capital in a relatively isolated place or to split the government across multiple capital cities that are quite distant from each other.[24] In particular, these examples show that including population concentration in our framework helps us account for countries that look like "outliers" with respect to the stylized fact linking inequality and population size.

We can also bring the evidence on the number of revolutions to bear on this issue. After all, even though our simple model rules out the occurrence of revolutions along the equilibrium path, one would still expect from our logic that, in practice, countries with larger capital cities would display a larger number of revolutionary episodes. To assess this conjecture, we collected data on the population of capital cities across countries, from the United Nations' World Urbanization Prospects, and we look at the impact of this variable on the measure of revolutions from Barro-Lee.[25] The results can be seen in Table 3, where we see a clear positive relationship: countries with larger capital cities do tend to have more revolutions, as our hypothesis would suggest. This relationship is much more robust than the one observed for total population size, as it is still significant even when including measures of openness and log land area – in fact, the effect of total population size vanishes once the population of the capital city is considered. Once again, the direct empirical evidence is consistent with our hypothesis concerning the relationship between population and revolutions, and it indeed indicates that taking the geographical distribution of the population into account is an important extension. [26]

[TABLE 3 HERE]

5 Empirical Evidence

5.1 Basic Results

The model proposed in the previous section has four basic predictions, all of which should hold *for non-democratic countries*:

- *Size effect*: There is a negative effect of population size on inequality in non-democratic countries.

  It follows that the stylized fact identified in Section 2 should be driven by non-democratic countries.

---

[24] In South Africa, Pretoria/Tshwane is the site of the executive, while the legislative and judicial branches are in Cape Town and Bloemfontein, respectively.

[25] We have data for 1990 and 2001, which we average. The results do not change if we use either year instead. Recall that the Barro-Lee data is for the period 1960-84, so we have to assume that the difference in periods is not too distortive. Since we are using cross-country variation, we feel this is not too bad an assumption.

[26] Ades and Glaeser (1995) also consider the empirical relationship between the number of revolutions and coups and the size of the main (not capital) city, but focusing on the effect of the former on the latter. They also obtain a positive coefficient. Their interpretation invites some caution in terms of interpreting causality, but the evidence is nevertheless consistent with our hypothesis.

- *The effect of density and concentration*: Our model implies that population density and concentration should also have a negative impact on inequality on top of the size effect.
- *Pre-tax vs. post-tax inequality*: Ours is a story about redistribution – population affects redistribution, which in turn affects observed inequality. As such, we would expect the size effect to appear mostly in post-tax inequality data.
- *Decreasing effect of population size:* The effect of size on redistribution should decrease as size itself increases, since the equilibrium redistribution increases at a diminishing rate.

We present some evidence that supports all four of these implications.

Democratic vs. non-democratic countries First, we split the sample according to whether $Polity\ 2 = 10$ or $Polity\ 2 < 10$ in the -10 to 10 scale of Polity IV. (The results are not sensitive to choosing other reasonable cutoffs, such as 9, 8 or 7.) The theory does not predict a relationship between population size and inequality in democratic countries, and Figure 4 (which like Figure 1 averages inequality over time and types of data) immediately suggests that indeed no decreasing pattern is present in that sub-sample. Table 4A confirms this result in the regression analysis. We repeat the basic specifications from Table 1 in Columns (1)-(4), and the coefficient of size in all specifications has a small size and is tested to be insignificantly different from zero.[27] Figure 4 suggests that results might be driven by the United States, which is distinctly larger and more unequal than the other OECD countries that constitute the bulk of the sample of democracies. Columns (5) and (6) confirm that this is not the case.[28]

[FIGURE 4 HERE]

[TABLE 4A HERE]

We then move to the analysis of non-democratic countries. In Figure 5 we again plot the raw correlation between the log of population and the Gini coefficient; this time there is some suggestion of the negative relationship predicted by the theory. The regression analysis in Table 4B shows that this is indeed the case: the coefficient on population size is robustly negative across all specifications and statistically significant whenever the relevant controls are included. Columns (2) and (3) successively

---

[27] We do not include the dummy variable for consumption-based inequality . It turns ou that very few of the (mainly OECD) countries in the sample have consumption-based data.

[28] Interestingly, we also find a large difference between gross and net income inequality, evidence of the high level of government transfers in democratic countries.

include the basic set of control variables.[29] Column (4) again shows that this result is robust to using clustered errors at the country level. Further checks of sensitivity are suggested by a look at Figure 5. First, it is apparent in that figure that many formerly socialist countries are grouped in a cluster of low inequality and relatively small population size; Column (5) includes legal origin dummies (which include a dummy for "socialist") to deal with such issues. The coefficient is still negative and significant. Column (6) shows a specification with regional dummies, in order to control for regional specificities, and the results are similar. Finally, Figure 5 again suggests the concern that the result may be driven by China and India; Column (7) shows that this is not the case.[30]

[FIGURE 5 HERE]

[TABLE 4B HERE]

The size of the relevant coefficient is quite stable across most specifications, once all the basic controls are included. The economic significance is even stronger than for the overall sample, as one would expect from our theory. In fact, the same comparison between countries at the 25th and 75th percentiles of the size distribution that we performed in Section 2 would now correspond to a shift of about three quarters of a standard deviation. Finally, it is also worth stressing that the results are similar, and similarly robust, when population in levels is used instead of its logarithm. (These and all other unreported results are available from the authors upon request.) The overall message is quite clear: our first prediction receives strong support from the data.

Population density While we leave the discussion of population concentration to subsection 5.3, we can provide evidence for our second prediction using data on population density. In fact, the prediction receives strong support from the raw scatterplot in Figure 6, where the Gini coefficients are plotted against a measure of density (population size divided by total land area) in logarithmic form.[31] Unlike Figure 5, this plot's negative slope is evidently not distorted by the presence of outliers. As anticipated in the previous section, including population density in the analysis helps us make sense of countries such as Brazil and South Africa, which would seem to challenge the connection between population and

---

[29] It is interesting to note the puzzling positive sign of the coefficient on democracy, suggesting that in non-democratic countries, a higher level of democracy is linked with higher, not lower inequality. This is not the case for the full sample, as democratic countries are typically less unequal than non-demoratic countries. While we leave this non-monotonicity as a topic for future research, it does reinforce the point suggested by our theory, that democracies and less-than-democratic countries are inherently different with respect to inequality and redistribution.

[30] Note that the dummy variable for gross vs. net income inequality is much smaller than that of democratic countries, making evident the lower level of government transfers compared to the latter. On the other hand, there is a large difference between consumption- and income-based inequalities, vindicating the inclusion of the dummy variable as suggested by Deininger and Squire(1996) and Atkinson and Brandolini(2001).

[31] We will discuss the evidence on population concentration separately, in the next subsection.

inequality. Table 5 shows the evidence from the regression analysis, by reproducing the exercise in Table 4B while including a measure of density. As suggested by the theory, the coefficient on density is negative and highly significant in all specifications, and the comparison with the population size coefficients shows that this finding is actually even more robust than the result on size itself. The coefficient on density is always highly statistically significant, even when errors are clustered at the country level and the three sets of dummy variables are included, which eventually renders population size itself insignificant (Columns (4)-(7)).

[FIGURE 6 HERE]

[TABLE 5 HERE]

Note that the analysis in Table 5 is restricted to the sample of non-democracies, to which all of our predictions pertain. While we do not report the results for the sample of democracies, in the interest of saving space, it is worth stressing that, just as in Table 4A, the relationship between inequality and population density also vanishes for those countries. Finally, we add that the results with levels instead of logs are still qualitatively similar. In sum, the empirical evidence strongly supports our theory's prediction that, in addition to the effect of population size, a non-democratic country that is more densely populated is likely to have a less unequal distribution.

Pre-tax vs. post-tax inequality Our third prediction has to do with the relationship between gross and net income inequality: our redistribution story leads us to expect the population effects in terms of size and density to be at work for the latter, but not for the former. While in the data it is the case that some redistribution policies impact pre-tax inequality – e.g. education policy – we would still expect the population effect to be stronger in post-tax inequality data. Table 6 supports this claim, by presenting two different kinds of test: including an interaction term between the dummy for gross inequality and population size, and splitting the sample along the gross versus net inequality dimension. In the former case, the prediction is that the population effect (a negative coefficient) is dampened for gross income inequality, and indeed, as displayed in Panel A (Columns (1)-(4)), the interaction term between gross-net separation and population size has a positive estimate, though not statistically significant. Column (4) also shows a positive effect of the interaction with the log of density, also consistent with the model's implication of an effect operating via redistribution. Panels B (Columns (5) and (6)) and C (Columns (7) and (8)) show the results when the sample is split (which leads to two subsamples of relatively similar size). They provide stronger evidence that in the "Gross" sample (Panel B) the effects are essentially absent, both for population size and density, when compared to the "Net" sample (Panel C). In other words, both exercises (especially the latter one) provide support for the contention that redistribution is the key for the negative relationship between population and inequality.

[TABLE 6 HERE]

Another way in which we investigate this prediction is to look at a tentative direct measure of redistribution, which is what our theory ultimately refers to. Since measures such as the size of government transfers are distorted by the fact that in many countries these transfers, to a large extent, are not in fact appropriated by the poor, we turn to a measure that uses the inequality data we have been analyzing: the difference between "Gross" and "Net" measures of inequality.[32] While this is an admittedly imperfect measure of the amount of redistribution that is effectively undertaken in a given country, it can nevertheless shed some extra light on the question at hand. After all, to the extent that an important part of redistribution policy takes part through taxes and transfers, our model would predict that this measure should display the positive effect of population on redistribution.

Since we are focusing on the different types of inequality, and also because data limitations mean that typically we do not have "Gross" and "Net" observation for a given country in a given year, we start by computing country averages of "Gross" and "Net" Gini coefficients over time. We had hitherto refrained from this kind of average, except in the context of scatterplots, precisely in order to keep clear the distinction between types of inequality. We first compare the results of the basic regression with the full set of controls (except for dummies, since the size of the sample is now fairly small) in the two samples. Columns (1) and (2) of Table 7 confirm the result from Table 6: the negative effect of population size and density is restricted to the "Net" sample.

[TABLE 7 HERE]

We then compute the difference between average "Gross" and "Net" Ginis, for the 39 countries that happen to have both types of measures. The results are in Columns (3)-(5) of Table 7. Using the same set of control variables as in Table 5, we see a very robust positive impact of density on redistribution, and a weaker but still present positive effect of population size (except in the last specification).[33] In sum, this evidence suggests that larger, more densely populated countries do redistribute more extensively. Put together with the results from Table 6, we have a picture where there is no population effect in the pre-tax

---

[32] Milanovic (2000) and Scruggs (2005) use a similar approach.

[33] Obviously clustering is not an issue here.

distribution of income, and a positive effect of population size and density on redistribution, leading to a negative effect on post-tax inequality. While we do stress that the imperfection of the measure requires some grains of salt in interpreting these results, they nevertheless add to the empirical support that our theory is able to muster. Incidentally, these results also speak to those obtained by Rodrik (1998). To the extent that his finding of a positive effect of openness (which displays a strong inverse correlation with size) on the size of government, through a mechanism involving a greater demand for redistribution as a means of insurance, could be extended to a prediction linking small size to greater demand for redistribution, our evidence seems to run in the opposite direction. In other words, our evidence suggests that there is no size effect on pre-tax inequality, but larger countries tend to redistribute more, not less, than smaller ones.

Diminishing effect of population While not as central to the logic of the theory as the previous three, since it relies on additional conditions on the distribution function, the prediction of a diminishing effect of population size on inequality provides another window into the interplay between these two variables. In order to test it, we split the sample into its four quartiles, and run the population size regressions, with the full set of non-dummy control variables, in each of the sub-samples. One observation is in order: unlike the previous three, this prediction refers to the second derivative of the relationship between population size and inequality, i.e. to the convexity of the functional form that governs such relationship. As the logarithm function is itself a concave function, using the logarithmic transformation of population may make this prediction undetectable. Therefore, we proceed with population in both logs and levels. Columns (1)-(4) of Table 8 show the results with logs, and, not surprisingly, the results are not very sharp. Columns (5)-(8) reveal a clearer pattern: the coefficient for the bottom quartile is an order of magnitude larger, in absolute value, than the coefficients for the middle quartiles; these are in turn an order of magnitude larger than the one for the top quartile. The fourth prediction of our theoretical framework thus also receives some empirical support.

[**TABLE 8 HERE**]

5.2 Robustness

Having presented the results in support of our four predictions, we can now assess their robustness along several dimensions.

Alternative measures of inequality Table 9 reports the main results using as the dependent variable (the negative of) the ratio between the share of income held by the bottom and the top quintiles, which also provides a measure of how equal the income distribution is. Columns (1)-(3) shows that the basic negative relationship between size and inequality also holds with this measure, and only for the sample of non-democracies. Columns (4) and (5) confirm the result with population density, and Columns (6)-(9) show that it is net inequality where the negative relationship is stronger, in terms of coefficient size and significance.

[TABLE 9 HERE]

Reverse causality and unobserved country-specific characteristics Similar to other cross-country empirical studies, the basic evidence presented so far could potentially suffer from endogeneity biases. On one hand, there might be the concern that the direction of causality could run from inequality to population; on the other hand, the existence of country-specific unobservables could lead to both larger population size and lower inequality. We can exploit the longitudinal structure of the data in order to allay these concerns with the use of dynamic panel techniques. Traditional panel methods such as random effects and fixed effects (within estimates) could suffer from serious endogeneity bias in a dynamic panel where present values of the dependent variable could affect future values of right-hand side variables. In addition, some variables in our analysis, such as population size and GDP, exhibit a high degree of persistence, as shown by standard unit root tests in panel. This destabilizes the dynamic panel technique proposed by Arellano and Bond (1991). We therefore opt for the two-step GMM approach proposed by Blundell and Bond (1998), using lagged differences as instruments for the equation in levels.[34]

The results with the Gini coefficient are presented in Table 10. Column (1) presents the basic specification, for non-democratic countries. Population size has a significantly negative impact on the Gini coefficient, controlling for the standard variables. In Column (2), we test the prediction on population density. This variable plays an important role, consistent with the theory, as it somewhat eclipses that of population size. Columns (3) to (5) present further sensitivity checks for these basic results. In Column (3), population and GDP are assumed to be endogenous (determined jointly at time t) as opposed to predetermined (determined at time $t-1$). This hypothesis removes the first lagged difference from the set of instruments; however, the qualitative results still obtain. Columns (4) and (5) show the results

---

[34] We use five lags of the main variables, namely population, density and GDP per capita, since more lags could lead to a problem of too many moment conditions. The application in Stata comes from Roodman (2006). Standard errors for the two-step procedure are corrected by Windmeijer's (2005) method.

also remain with different sets of additional controls.

[TABLE 10 HERE]

Columns (6)-(8) explore in greater detail our additional predictions. Column (6) compares democratic vs. non-democratic countries, confirming the positive estimates of interactions of the full democracy dummy with population size and density. Columns (7) and (8) in turn compare gross and net income inequalities. The interaction between "Gross" and population is significantly positive and sizeable, showing that the relationship between population and inequality predicted by our model effectively occurs mostly for net income inequality. Dynamic panel specification tests are satisfied thoroughly: the Hansen-Sargan overidentification test for each set of instruments is always passed, and the Arellano-Bond tests of serial autocorrelation of first-differenced residuals also show evidence of first-order autocorrelation, while none at the second order.

In sum, Table 10 shows strong and robust evidence in support of the predicted relationships when country-specific unobservables are duly taken into account.

5.3 Population Concentration

As anticipated in subsection 4.2.1, we would like to have a measure of population concentration around the capital city, in order to test more precisely the second prediction stemming from our paper. Since no such measure is readily available, we develop our own measure of population corrected for concentration, which we term “geo-populational potential”. (A full description and discussion of this measure and its axiomatic foundation can be found in Campante and Do 2007.) This measure captures the total population, but with each individual weighted by the inverse of the logarithm of her distance to the capital city. Formally, this potential is defined as:

$$\mathbf{GPP} = \int_0^{\overline{\delta}} g(\delta)d\delta - \frac{1}{\log(\overline{\delta})}\int_0^{\overline{\delta}} \log_+ \delta g(\delta)d\delta, \tag{10}$$

and if normalized by the population size, $\int_0^{\overline{\delta}} g(\delta)d\delta$, it becomes the index of population concentration:

$$\mathbf{PopCon} = 1 - \frac{1}{\log(\overline{\delta})\int_0^{\overline{\delta}} g(\delta)d\delta}\int_0^{\overline{\delta}} \log_+ \delta g(\delta)d\delta,$$

In this expression, $\overline{\delta}$, the maximum possible value of δ, is used to normalize the value of **PopCon** so that its range is from 0 (representing a situation in which the entire population lives at a distance $\overline{\delta}$ from the empty capital city) to 1 (when the entire population lives in the capital).[35]

---

[35] Note that $\log_+ \delta$ is defined as $max\{\log(\delta), 0\}$, a normalization that takes care of possible negative values of the log of distance.

We construct these measures for a cross-country sample for the year 1990, using Columbia University's Gridded Population of the World (GPW) dataset version 3.0 (CIESIN et al., 2004), arguably the most up-to-date global map of population distribution. This dataset provides population measures for each 2.5 arc-minute cell (around 85 squared kilometers), which is taken to be a population mass point at the center of the cell for the calculation of the **PopCon** and **GPP** indices.[36] We end up with a sample of **PopCon** ranging from 0.245 (United States) to 0.989 (Monaco), with a mean of 0.511 and a standard deviation of 0.132. It is worth noting that the lower end of the range indeed displays countries like Brazil (2nd lowest at 0.247), and South Africa (4th lowest at 0.263), which have already appeared in our discussion on the location of capital cities, and which might have been considered to be counter-examples to the negative relationship between inequality and population size.

We are thus equipped to test the model's predictions on population concentration. Columns (1) to (4) in Table 11, show the OLS regressions of the country Gini coefficients (computed as averages over the available sample) on population and population concentration measures, plus assorted controls, for the sample of non-democratic countries. The first two columns suggest that population concentration has indeed a significant negative impact on inequality, beyond the impact of population size, very much in accordance with our theory. A coefficient of *log* **PopCon** of size -9.5 means that an increase of a standard deviation in *log* **PopCon** leads to a quite sizeable decrease of 2.2 Gini points; and a doubled **PopCon** implies a decrease of 6.6 Gini points.

[TABLE 11 HERE]

We then divide the sample into "Gross" and "Net" subsamples, in Columns (3) and (4). While most coefficients of interest are not statistically significant, due to the small sample size, we can see a large difference between the two columns regarding both the coefficient of population size and that of population concentration. The larger coefficient from the "Net" sample supports the predictions that the impacts of population size and population concentration work mostly through redistribution.

The use of population concentration may give rise to questions on its reverse causality with respect to inequality. For a given population size, higher inequality could induce the population to disperse in search of, for instance, better local political influence, thereby diminishing population concentration and creating a negative relationship between inequality and population concentration that is fundamentally different from our theory. In order to correct for this problem, we use population density as an instrumental

---

[36] Details of the calculation of these indices can also be found in Campante and Do (2007).

variable for population concentration. The results are presented in Columns (5) to (8).[37] Columns (5) and (6) confirm the pattern found in the OLS regressions of Columns (1) and (2), with coefficients of *log***PopCon** close to the OLS findings. Columns (7) and (8) also confirm the general pattern found in Columns (3) and (4).

In sum, our predictions regarding the effect of population concentration receive robust support from the data.

6 Alternative Explanations

The breadth of empirical evidence compiled in the previous section clearly suggests that our explanation based on the political economy of redistribution goes some way towards explaining the empirical regularity we uncovered, linking inequality and population. Another way to assess the validity of that particular explanation is to consider alternative factors, apart from our political economy considerations, that could generate the basic regularity, and see if they conform to the evidence we present. We do so at a brief and informal level, just so that we are able to contrast the main implications such factors would likely involve, and how they relate to the data.

For instance, one could imagine that the size effect is engendered by the existence of some fixed cost of setting up a redistribution system: If that is the case, countries with larger populations would be better able to dilute that cost, and would thus be more effective at redistributing. This would be clearly consistent with the evidence that the size effect relates to redistribution and post-tax inequality, and perhaps such an explanation could be extended to include the prediction about density and population concentration – say, by imagining that fixed cost to be location-specific. But it is far from clear how such an explanation would generate a different behavior in democracies and non-democracies: Why would such fixed costs of redistribution be present in the latter, but not in the former?

The fixed cost intuition could perhaps be applied to a context other than that of setting up a redistributive system. For instance, it could be the case that the fixed cost could be related to the building of infrastructure, the presence of which could in turn lead to lower inequality – perhaps by reducing inequality between regions. Once again, it is far from clear that this argument would be able to generate the distinction between democracies and non-democracies. In addition, such a story would not be consistent with the fact that the size effect is verified for post-tax, but not pre-tax inequality.

Moving away from the idea of fixed costs, another interpretation could be related to diversification:

---

[37] All first stage regressions have high F-statistics and present no problem of weak instruments; they are available upon request.

Perhaps larger countries are able to sustain more diversified economies, and this may in turn lead to more opportunities for insurance against idiosyncratic shocks, and thus less inequality. This would be hard to reconcile with the empirical finding that greater population density and concentration are even more strongly related with lower inequality than population size itself: It is not clear how increased population density or concentration would be linked to greater diversification. In addition, such an explanation does not seem consistent with the fact that the empirical regularity seems to be related to redistribution, and not to pre-tax inequality.

While this list is obviously not exhaustive, the fact that the empirical evidence helps us refute some of the most natural alternative explanations increases our confidence that the mechanism we propose is important in accounting for the stylized fact under examination.

7 Concluding Remarks

While we have started by uncovering a puzzling empirical regularity – a negative relationship between population size and inequality –, our theory enabled us to rephrase this observation: *there is a negative relationship between population size, and its geographic concentration, and post-tax inequality, in non-democratic countries.* Moreover, the empirical support that we obtained for this proposition suggests that the mechanism the theory advances, having to do with the way redistributive demands are expressed when democratic channels are blocked and how this is affected by population size, has good explanatory power.

This exercise sheds light on the determinants of inequality and redistribution, and an interesting general point can be taken from it: there is an important difference in how democratic and non-democratic polities deal with these issues. While this is hardly a surprising observation, and has in fact been explored in the literature before (e.g. Persson and Tabellini 1994), one contribution of our model is to illustrate how systematic patterns with respect to other variables of economic interest may vary along those lines, as democracies and non-democracies give different weights to different forms of political activity, which react in distinct manners to such variables. This point is broader than the specific stylized contrast between "electoral" and "revolutionary" channels, and it underscores the idea that it may be ill-advised to extrapolate evidence from one specific set of countries to other very dissimilar ones.

This paper also opens up some clear avenues for future research. On the theory side, one could push further the theory of revolutions that is sketched here. In particular, the issues of coordination that are touched upon here merit further consideration. As far as empirics are concerned, a natural extension would be to look for micro-level data. Data within a country, across administrative zones, presumably represent a much more homogenous sample, and as a result omitted variable bias is much less of a problem. When they are collected, micro-level panel data are usually more balanced over time, in terms of availability, quality and compatibility. This would bring another dimension to the test of the predictions.

We are currently pursuing such an approach, in joint work with our colleague Li Han, using a household-level dataset from 48 randomly selected villages in China. These data have information on taxes set by different levels of administration, which enables us to compute a direct measure of tax-induced redistribution at the village level. We also have information from a special natural experiment on grassroots politics in China, in which direct elections for village committees have been introduced. This provides us with variation in the level of "democratic" channels available. Since we could safely assume that village politics in China is at best imperfectly democratic, our theory predicts that redistribution rises with population size. Moreover, in non-democratic villages that effect is more important than in those having started their way into democracy, i.e. those holding elections with multiple candidates. These are two main predictions that we can eventually test in this context, and some very preliminary evidence shows some support for them. We hope to have results from this exercise in the near future.

References

[1] Acemoglu, Daron and James Robinson (2000), “Why Did the West Extend the Franchise?: Democracy, Inequality and Growth in Historical Perspective”, *Quarterly Journal of Economics* 115: 1167-1199.

[2] Acemoglu, Daron and James Robinson (2005), *Economic Origins of Dictatorship and Democracy.* Cambridge, UK: Cambridge University Press.

[3] Ades, Alberto F. and Edward L. Glaeser (1995), “Trade and Circuses: Explaining Urban Giants”, *Quarterly Journal of Economics* 110: 228-258.

[4] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg (2003), “Fractionalization,” *Journal of Economic Growth* 8: 155-194.

[5] Alesina, Alberto and Edward L. Glaeser (2004), *Fighting Poverty in the US and Europe: A World of Difference.* Oxford, UK: Oxford University Press.

[6] Alesina, Alberto and Eliana La Ferrara (2005), “Preferences for Redistribution in the Land of Opportunities”, *Journal of Public Economics* 89: 897-931.

[7] Alesina, Alberto and Enrico Spolaore (2003), *The Size of Nations.* Cambridge, MA: MIT Press.

[8] Alesina, Alberto and Romain Wacziarg (1998), “Openness, Country Size and Government”, *Journal of Public Economics* 69: 305-321.

[9] Arellano, Manuel and Stephen Bond (1991), “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations”, *The Review of Economic Studies* 58: 277-97.

[10] Arellano, Manuel and Bo Honoré (2001), “Panel Data Models: Some Recent Developments”, in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Vol. 5, Ch. 53.

[11] Atkinson, Anthony B. and Andrea Brandolini (2001), “Promise and Pitfalls of Using ‘Secondary’ Data-Sets: Income Inequality in OECD Countries as a Case Study”, *Journal of Economic Literature*, 39:771-799.

[12] Banerjee, Abhijit and Esther Duflo (2003), ““Inequality and Growth: What Can the Data Say?”, *Journal of Economic Growth* 8: 267-299.

[13] Bénabou, Roland (2000), “Unequal Societies: Income Distribution and the Social Contract”, *American Economic Review* 90: 96-129.

[14] Bénabou, Roland and Efe Ok (2001), “Social Mobility and the Demand for Redistribution: The POUM Hypothesis”, *Quarterly Journal of Economics* 116: 447-487.

[15] Blundell, Richard and Stephen Bond (1998), “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”, *Journal of Econometrics* 87: 115-43.

[16] Bolton, Patrick and Gerard Roland (1997), “The Breakup of Nations: A Political Economy Analysis”, *Quarterly Journal of Economics* 112: 1057-1090.

[17] Bourguignon, François (1979), ”Decomposable Income Inequality Measures”, *Econometrica* 47: 901-920.

[18] Brautigam, Deborah and Michael Woolcock (2001), “Small States in a Global Economy: The Role of Institutions in Managing Vulnerability and Opportunity in Small Developing Countries”, WIDER Discussion Paper #2001/37, Helsinki: UNU/WIDER.

[19] Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow (2003), *The Logic of Political Survival.* Cambridge, MA: MIT Press.

[20] Campante, Filipe R. and Quoc-Anh Do (2007), “Axiomatic Foundation For An Index Of Population Concentration”, Harvard University (mimeo)

[21] Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT), (2004), “Gridded Population of the World (GPW), Version 3”, Columbia University. Available at http://beta.sedac.ciesin.columbia.edu/gpw.

[22] Couto, Ronaldo Costa (2001), *Brasília Kubitschek de Oliveira.* São Paulo: Record.

[23] Deininger, Klaus and Lyn Squire (1996), “A New Data Set Measuring Income Inequality”, *The World Bank Economic Review* 10: 565-591.

[24] Engerman, Stanley and Kenneth Sokoloff (2002), “Factor Endowments, Inequality, and Paths of Development among New World Economies”, *Economía* 3: 41-109.

[25] Fearon, James D. and David D. Laitin (2003), ”Ethnicity, Insurgency, and Civil War”, *American Political Science Review* 97: 75-90.

[26] Granovetter, Mark (1978), “Threshold Models of Collective Behavior”, *American Journal of Sociology* 83: 1420-1443.

[27] Grossman, Herschel I. (1995), “Robin Hood and the Redistribution of Property Income”, *European Journal of Political Economy* 11: 399-410.

[28] Grossman, Herschel I. and Murat F. Iyigun (1997), “Population Increase and the End of Colonialism”, *Economica* 64: 483-493.

[29] *International Herald Tribune* (2005) “A Government on a Move to a Half-Built Capital”, November 11th issue.

[30] Katzenstein, Peter (1985), *Small States in World Markets: Industrial Policy in Europe.* Ithaca, NY: Cornell University Press.

[31] Kuran, Timur (1989), “Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution”, *Public Choice* 61: 41-74.

[32] Kuran, Timur (1995), “The Inevitability of Future Revolutionary Surprises”, *American Journal of Sociology* 100:1528-1551.

[33] Le Bon, Gustave (1913), *The Psychology of Revolution* (transl. by Bernard Miall). New York: G. P. Putnam’s Sons.

[34] Lohmann, Susanne (1994), “Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991”, *World Politics* 47: 42-101.

[35] Merrill, Tim (ed.) (1994), *Nicaragua: A Country Study.* Washington DC: Federal Research Division, Library of Congress. (Available at http://lcweb2.loc.gov/frd/cs/nitoc.html)

[36] Milanovic, Branko (2000), “The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data”, *European Journal of Political Economy* 16: 367-410.

[37] Mulligan, Casey B. and Andrei Shleifer (2005), “The Extent of the Market and the Supply of Regulation”, *Quarterly Journal of Economics* 120: 1445-1473.

[38] Persson, Torsten and Guido Tabellini (1994), “Is Inequality Harmful for Growth?”, *American Economic Review* 84: 600-621.

[39] Rodrik, Dani (1998), “Why Do More Open Economies Have Bigger Governments?”, *Journal of Political Economy* 106: 997-1032.

[40] Roemer, John E. (1985), ”Rationalizing Revolutionary Ideology”, *Econometrica* 53: 85-108.

[41] Roodman, David M. (2006), “How to Do xtabond2: An Introduction to ‘Difference’ and ‘System’ GMM in Stata”, Center for Global Development Working Paper 103, Washington DC.

[42] Rose, Andrew K. (2006), “Size Really Doesn’t Matter: In Search of a National Scale Effect”, NBER WP 12191.

[43] Scruggs, Lyle (2005), “Redistributive Consequences of Welfare State Entitlements”, University of Connecticut (mimeo).

[44] Spielvogel, Jackson (1999) *Western Civilization: Comprehensive Volume*, Belmont, CA: Wadsworth Publishing.

[45] *The Economist* (1997), “Capital Punishments,” December 17th issue.

[46] *The Economist* (2005), “Another Autocrat Falls,” March 31st issue.

[47] *The Economist* (2006), “Waving the Denim”, March 18th issue.

[48] *The Economist* (2006), “Old Soldiers, Old Habits”, September 21st issue.

[49] Trotsky, Leon (1932), *The History of The Russian Revolution.* (transl. by Max Eastman), New York: Simon and Schuster.

[50] Windmeijer, Frank (2005), “A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators”, *Journal of Econometrics* 126: 25-51.

[51] World Institute of Development Economic Research (2005), ”World Income Inequality Database v. 2.0a”, available at http://www.wider.unu.edu/wiid/wiid.htm.

[52] Yitzhaki, Shlomo (1996), “On Using Linear Regression in Welfare Economics”, *Journal of Business and Economic Statistics*, 14: 478-486.

8 Appendix

Here we provide the details for the case in which we allow $\overline{\pi}$ to vary with N. In that case, we can write:

$$\tau_v^P = \tau_m - \frac{K/\overline{\pi}(N) - b_1}{\bar{w} - \Omega} \tag{11}$$

Computing the derivative of this with respect to N yields:

$$\frac{\partial \tau_v^P}{\partial N} = \frac{K/\overline{\pi}(N) - b_1}{(\bar{w} - \Omega)^2} \frac{1}{F'(\Omega)} \frac{\overline{n}_r}{N^2} + \frac{\left(K/\overline{\pi}(N)^2\right) \overline{\pi}'(N)}{\bar{w} - \Omega} \tag{12}$$

The first term is positive, as shown in Proposition 1. The second term will also be positive, reinforcing the result, if $\overline{\pi}'(N) > 0$, meaning that there are diseconomies of scale in counter-revolution: it is harder to fight revolution attempts in larger countries. If the opposite is true, then the condition for the overall effect to be positive is:

$$\left(K/\overline{\pi}(N)^2\right) \overline{\pi}'(N) \left(\bar{w} - \Omega\right) F'(\Omega) N^2 + \left(K/\overline{\pi}(N) - b_1\right) \overline{n}_r > 0$$

In general, this is true if the economies of scale are not strong enough:

$$|\overline{\pi}'(N)| < \frac{\left(K/\overline{\pi}(N) - b_1\right) \overline{n}_r \overline{\pi}(N)^2}{K \left(\bar{w} - \Omega\right) F'(\Omega) N^2}$$

More specifically, when N is small, the first term in (12) dominates, and the effect is positive, provided that we assume that $\overline{\pi}'(N)$ does not grow without bound (or at least not fast enough) as $N \to 0$. Similarly, when N is large, we can get a similar result if we impose that $\overline{\pi}(N)$ is bounded away from zero, that $\lim_{N\to\infty} F'(N) > 0$, and that $\lim_{N\to\infty} N^2 |\overline{\pi}'(N)| = 0$, meaning that the economies of scale decrease fast enough.

Table 1
Inequality and Population Size

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Population	**0.153**	**-1.115**	**-2.23**	**-2.23**	**-2.48**	**-0.924**	**-1.547**
	[0.323]	**[0.407]*** **	**[0.418]*** **	**[0.655]*** **	**[0.650]*** **	**[0.640]**	**[0.647]** **
Consumption dummy	1.23	-4.474	-5.228	-5.228	-5.234	-4.106	-5.304
	[1.386]	[1.330]***	[1.440]***	[1.377]***	[1.243]***	[1.194]***	[1.377]***
Gross dummy	2.239	2.072	1.935	1.935	2.767	3.384	1.795
	[1.137]**	[1.054]**	[0.988]*	[1.225]	[1.050]***	[1.017]***	[1.224]
Log GDP per capita		-4.969	-4.623	-4.623	-5.375	-4.064	-4.939
		[0.476]***	[0.593]***	[0.751]***	[0.819]***	[0.939]***	[0.759]***
Openness		-0.05	-0.043	-0.043	-0.011	-0.005	-0.034
		[0.017]***	[0.015]***	[0.019]**	[0.019]	[0.019]	[0.019]*
ELF in 1985			9.332	9.332	1.097	5.21	9.413
			[2.010]***	[2.944]***	[3.192]	[3.013]*	[3.025]***
Polity2			0.268	0.268	0.18	0.119	0.293
			[0.096]***	[0.128]**	[0.119]	[0.121]	[0.133]**
Log Land Area			1.373	1.373	1.774	0.804	1.318
			[0.369]***	[0.558]**	[0.558]***	[0.447]*	[0.548]**
Legal Origin dummies					Yes		
Regional dummies						Yes	
China & India dummies							Yes
Clustered Errors				Yes	Yes	Yes	Yes
Observations	1352	1116	968	968	968	968	968
R-squared	0.009	0.261	0.389	0.389	0.539	0.598	0.41

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Intercepts are omitted. Clustered errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2
Revolutions and Population Size

Dependent variable: Average Number of Revolutions	(1)	(2)	(3)	(4)	(5)	(6)
Log Population	**0.022**	**0.018**	**0.018**	**0.021**	**0.01**	**-0.004**
	[0.010]**	**[0.010]***	**[0.010]***	**[0.012]***	**[0.020]**	**[0.025]**
Log GDP per capita		-0.063	-0.064	-0.113	-0.115	-0.108
		[0.020]***	[0.019]***	[0.037]***	[0.037]***	[0.039]***
ELF in 1985			0.045	0.006	-0.011	-0.004
			[0.074]	[0.060]	[0.063]	[0.062]
Polity2				0.012	0.012	0.012
				[0.006]*	[0.006]*	[0.007]*
Log Land Area					0.013	0.011
					[0.018]	[0.019]
Openness						-0.001
						[0.001]
Observations	79	77	76	74	74	73
R-squared	0.033	0.15	0.172	0.229	0.236	0.249

Robust standard errors in brackets. Intercepts are omitted.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3
Revolutions and Size of Capital Cities

Dependent variable: Average Number of Revolutions	(1)	(2)	(3)	(4)	(5)	(6)
Log Population of Capital	**0.032**	**0.052**	**0.055**	**0.07**	**0.072**	**0.069**
	[0.024]	**[0.023]****	**[0.027]****	**[0.027]****	**[0.028]****	**[0.027]****
Log Population	0.002	-0.013	-0.016	-0.02	-0.034	-0.041
	[0.018]	[0.017]	[0.020]	[0.018]	[0.029]	[0.032]
Log GDP per capita		-0.072	-0.065	-0.125	-0.127	-0.119
		[0.019]***	[0.018]***	[0.036]***	[0.036]***	[0.039]***
ELF in 1985			0.098	0.068	0.049	0.057
			[0.081]	[0.062]	[0.059]	[0.060]
Polity2				0.015	0.015	0.014
				[0.006]**	[0.006]**	[0.007]**
Log Land Area					0.015	0.011
					[0.017]	[0.018]
Openness						-0.001
						[0.001]
Observations	72	71	70	69	69	68
R-squared	0.063	0.211	0.234	0.276	0.276	0.276

Robust standard errors in brackets. Intercepts are omitted.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4A
Inequality and Population Size: Democratic Countries
(Polity2 = 10)

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)
Log Population	**0.407**	**0.177**	**-0.087**	**-0.087**	**-0.478**	**-0.478**
	[0.381]	**[0.423]**	**[0.440]**	**[0.669]**	**[0.508]**	**[0.684]**
Gross dummy	5.476	5.551	5.879	5.879	5.832	5.832
	[0.875]***	[0.990]***	[0.950]***	[1.358]***	[0.960]***	[1.345]***
Log GDP per capita		-4.932	-5.072	-5.072	-5.107	-5.107
		[1.114]***	[1.073]***	[1.485]***	[1.060]***	[1.431]***
Openness		-0.026	-0.03	-0.03	-0.038	-0.038
		[0.017]	[0.017]*	[0.027]	[0.017]**	[0.027]
ELF in 1985			8.05	8.05	7.048	7.048
			[2.377]***	[3.900]**	[2.460]***	[3.860]*
Log Land Area			0.06	0.06	-0.147	-0.147
			[0.332]	[0.654]	[0.335]	[0.639]
US dummy					6.968	6.968
					[1.516]***	[2.254]***
Clustered Errors				Yes		Yes
Observations	613	528	470	470	470	470
R-squared	0.13	0.296	0.39	0.39	0.427	0.427

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Constants are omitted. Clustered errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4B
Inequality and Population Size: Non-Democratic Countries
(Polity2 < 10)

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Population	**-0.517**	**-1.712**	**-3.575**	**-3.575**	**-3.041**	**-2.169**	**-2.731**
	[0.356]	**[0.487]*** **	**[0.559]*** **	**[0.816]*** **	**[0.873]*** **	**[0.908]** **	**[0.855]*** **
Consumption dummy	-3.631	-5.973	-6.572	-6.572	-6.567	-5.443	-6.836
	[1.584]**	[1.562]***	[1.713]***	[1.604]***	[1.264]***	[1.285]***	[1.618]***
Gross dummy	1.035	1.044	0.01	0.01	0.446	1.376	-0.331
	[1.539]	[1.501]	[1.556]	[1.798]	[1.442]	[1.447]	[1.804]
Log GDP per capita		-3.953	-3.161	-3.161	-3.377	-3.673	-3.472
		[0.763]***	[0.883]***	[1.077]***	[1.026]***	[1.227]***	[1.073]***
Openness		-0.067	-0.064	-0.064	0	-0.009	-0.055
		[0.025]***	[0.018]***	[0.024]***	[0.021]	[0.023]	[0.024]**
ELF in 1985			9.218	9.218	1.474	4.363	9.527
			[2.592]***	[3.812]**	[3.950]	[3.629]	[3.961]**
Polity2			0.298	0.298	0.167	0.151	0.311
			[0.094]***	[0.120]**	[0.108]	[0.112]	[0.128]**
Log land area			2.365	2.365	2.085	1.336	2.247
			[0.495]***	[0.678]***	[0.768]***	[0.646]**	[0.659]***
Legal Origin dummies					Yes		
Regional dummies						Yes	
China & India dummies							Yes
Clustered Errors				Yes	Yes	Yes	Yes
Observations	642	564	498	498	498	498	498
R-squared	0.039	0.196	0.369	0.369	0.56	0.588	0.39

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Intercepts omitted. Clustered errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5
Inequality and Population Density in Non-Democratic Countries

(Polity2 < 10)

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Population	**0.14**	**-1.18**	**-1.209**	**-1.209**	**-1.164**	**-0.833**	**-0.484**
	[0.366]	**[0.465]****	**[0.490]****	**[0.735]**	**[0.723]**	**[0.787]**	**[0.819]**
Log Density	**-3.309**	**-3.158**	**-2.365**	**-2.365**	**-2.377**	**-1.337**	**-2.247**
	[0.560]***	**[0.566]*****	**[0.496]*****	**[0.678]*****	**[0.677]*****	**[0.646]****	**[0.659]*****
Consumption dummy	-3.837	-5.996	-6.572	-6.572	-6.56	-5.444	-6.836
	[1.456]***	[1.448]***	[1.713]***	[1.604]***	[1.608]***	[1.285]***	[1.618]***
Gross dummy	0.967	0.994	0.01	0.01	-0.039	1.376	-0.331
	[1.445]	[1.410]	[1.556]	[1.798]	[1.793]	[1.447]	[1.804]
Log GDP per capita		-3.69	-3.162	-3.162	-3.172	-3.674	-3.472
		[0.724]***	[0.883]***	[1.077]***	[1.075]***	[1.227]***	[1.073]***
Openness		-0.07	-0.064	-0.064	-0.066	-0.009	-0.055
		[0.018]***	[0.018]***	[0.024]***	[0.024]***	[0.023]	[0.024]**
ELF in 1985			9.218	9.218	8.649	4.363	9.528
			[2.592]***	[3.812]**	[3.830]**	[3.629]	[3.961]**
Polity2			0.298	0.298	0.299	0.151	0.311
			[0.094]***	[0.120]**	[0.120]**	[0.112]	[0.128]**
Legal Origin dummies					Yes		
Regional dummies						Yes	
China & India dummies							Yes
Clustered Errors				Yes	Yes	Yes	Yes
Observations	638	558	498	498	498	498	498
R-squared	0.138	0.224	0.331	0.331	0.344	0.619	0.357

Weighted Least Squares regressions, weight = (number of observations per country)$^{-1}$. Robust standard errors in brackets. Intercepts are omitted. Clustered errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6
Inequality and Population in Non-Democratic Countries: Gross versus Net Inequality
(Polity2 < 10)

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Full Sample				Panel B: "Gross" Sample		Panel C: "Net" Sample	
Log Population	-1.606	-1.606	-1.553	-1.553	**-0.289**	**-0.289**	**-1.827**	**-1.827**
	[0.625]**	[0.876]*	[0.654]**	[0.944]	**[0.791]**	**[0.898]**	**[0.701]*** **	**[1.039]***
Log Density	-2.324	-2.324	-2.487	-2.487	**-1.419**	**-1.419**	**-2.667**	**-2.667**
	[0.531]***	[0.710]***	[0.579]***	[0.818]***	**[0.882]**	**[1.042]**	**[0.693]*** **	**[0.920]*** **
Consumption dummy	-6.329	-6.329	-6.304	-6.304			-6.825	-6.825
	[1.681]***	[1.552]***	[1.698]***	[1.577]***			[1.787]***	[1.737]***
Gross dummy	-18.077	-18.077	-18.668	-18.668				
	[13.216]	[15.978]	[13.575]	[16.004]				
Log GDP per capita	-3.162	-3.162	-3.149	-3.149	-1.117	-1.117	-4.631	-4.631
	[0.884]***	[1.078]***	[0.889]***	[1.082]***	[1.240]	[1.400]	[1.105]***	[1.307]***
Openness	-0.063	-0.063	-0.062	-0.062	-0.05	-0.05	-0.076	-0.076
	[0.020]***	[0.025]**	[0.020]***	[0.025]**	[0.035]	[0.037]	[0.023]***	[0.032]**
ELF in 1985	9.012	9.012	9.059	9.059	18.445	18.445	4.152	4.152
	[2.599]***	[3.770]**	[2.624]***	[3.804]**	[3.369]***	[4.409]***	[3.386]	[4.499]
Polity2	0.303	0.303	0.302	0.302	0.459	0.459	0.225	0.225
	[0.094]***	[0.119]**	[0.093]***	[0.117]**	[0.176]***	[0.207]**	[0.098]**	[0.122]*
Log Pop X Gross	**1.106**	**1.106**	**1.048**	**1.048**				
	[0.779]	**[0.943]**	**[0.793]**	**[0.991]**				
Log Density X Gross			**0.375**	**0.375**				
			[1.106]	**[1.328]**				
Clustered Errors		Yes		Yes		Yes		Yes
Observations	498	498	498	498	204	204	294	294
R-squared	0.375	0.375	0.375	0.375	0.358	0.358	0.403	0.403

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Intercepts are omitted. Clustered errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7
Redistribution and Population in Non-Democratic Countries
(Polity2<10)

	(1)	(2)	(3)	(4)	(5)
	“Gross” Gini	“Net” Gini	“Gross” Gini – “Net” Gini		
Log Population	**-0.204**	**-2.15**	**0.868**	**0.261**	**-0.216**
	[1.111]	**[1.100]***	**[0.859]**	**[1.237]**	**[1.151]**
Log Density	**-0.81**	**-3.024**	**2.89**	**3.034**	**4.585**
	[1.039]	**[1.057]*****	**[1.099]****	**[1.130]****	**[0.920]*****
Log GDP per capita	-1.201	-3.663		0.101	2.148
	[1.640]	[1.678]**		[1.548]	[1.981]
Openness	-0.045	-0.081		-0.031	-0.049
	[0.050]	[0.055]		[0.049]	[0.049]
ELF in 1985	20.848	2.425			17.666
	[4.835]***	[4.786]			[5.343]***
Polity2	0.689	0.321			0.446
	[0.273]**	[0.198]			[0.348]
Observations	51	61	39	39	39
R-squared	0.421	0.364	0.16	0.167	0.428

* significant at 10%; **significant at 5%; *** significant at 1%

Table 8
Inequality and Population in Non-Democratic Countries: Convexity

(Polity2 < 10)

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Log Population	**-5.275**	**-2.63**	**-7.687**	**-6.252**				
	[2.066]**	**[2.465]**	**[5.167]**	**[1.100]*****				
Population Level					**-1.89E-06**	**-2.48E-07**	**-2.56E-07**	**-1.60E-08**
					[8.790e-07]**	**[2.119e-07]**	**[1.511e-07]***	**[2.251e-09]*****
Consumption dummy	-6.59	-9.866	0.051	-1.124	-6.788	-9.81	-0.285	-2.288
	[2.404]***	[2.007]***	[2.276]	[3.064]	[2.437]***	[1.987]***	[2.148]	[3.025]
Gross dummy	-4.222	0.173	7.183	5.383	-3.715	0.253	7.011	4.143
	[3.028]	[2.044]	[3.567]**	[2.307]**	[2.980]	[1.993]	[3.300]**	[2.208]*
Log GDP per capita	-7.515	-1.237	1.789	0.351	-7.43	-1.266	1.8	0.464
	[1.801]***	[1.414]	[2.043]	[1.281]	[1.806]***	[1.372]	[2.079]	[1.284]
Openness	-0.004	-0.119	-0.019	0.014	0.002	-0.117	-0.014	0.027
	[0.032]	[0.031]***	[0.031]	[0.026]	[0.034]	[0.031]***	[0.031]	[0.026]
ELF in 1985	-3.707	19.164	17.771	2.625	-2.703	19.527	17.995	2.375
	[7.629]	[6.726]***	[4.608]***	[2.610]	[7.628]	[6.713]***	[4.326]***	[2.590]
Polity2	0.166	0.436	-0.222	0.256	0.165	0.442	-0.21	0.22
	[0.204]	[0.134]***	[0.191]	[0.137]*	[0.212]	[0.134]***	[0.193]	[0.142]
Log Land Area	-0.118	0.296	4.015	2.983	-0.197	0.197	3.992	2.28
	[0.844]	[1.438]	[0.995]***	[0.689]***	[0.852]	[1.466]	[0.982]***	[0.587]***
Observations	138	116	109	141	138	116	109	141
R-squared	0.352	0.689	0.568	0.571	0.342	0.69	0.581	0.568

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Constants are omitted.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9
Inequality and Population: Alternative Measure of Inequality

Dependent variable: - Q1/Q5*100%	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Polity2=10	Polity2<10							
		Full sample				Gross sample		Net sample	
Log Population	**-0.749**	**-2.182**	**-2.182**	**-0.572**	**-0.572**	**-0.407**	**-0.407**	**-0.703**	**-0.703**
	[0.476]	**[0.390]*** **	**[0.628]*** **	**[0.400]**	**[0.597]**	**[0.620]**	**[0.685]**	**[0.589]**	**[0.817]**
Log Density				**-1.61**	**-1.61**	**-1.166**	**-1.166**	**-1.715**	**-1.715**
				[0.368]* **	**[0.498]*** **	**[0.723]**	**[0.841]**	**[0.442]*** **	**[0.584]*** **
Consumption dummy	1.31	-6.912	-6.912	-6.912	-6.912			-6.553	-6.553
	[1.928]	[1.312]***	[1.178]***	[1.312]***	[1.178]***			[1.380]***	[1.279]***
Gross dummy	6.26	-3.165	-3.165	-3.165	-3.165				
	[1.206]***	[1.637]*	[1.623]*	[1.637]*	[1.623]*				
Log GDP per capita	-3.801	-2.563	-2.563	-2.564	-2.564	-2.306	-2.306	-2.602	-2.602
	[1.690]**	[0.701]***	[0.899]***	[0.701]***	[0.899]***	[1.205]*	[1.440]	[0.799]***	[1.029]**
Openness	-0.038	-0.028	-0.028	-0.028	-0.028	-0.022	-0.022	-0.032	-0.032
	[0.020]*	[0.013]**	[0.017]	[0.013]**	[0.017]	[0.020]	[0.024]	[0.019]*	[0.025]
ELF in 1985	2.673	4.861	4.861	4.861	4.861	11.801	11.801	2.094	2.094
	[2.676]	[2.183]**	[3.273]	[2.183]**	[3.273]	[3.825]***	[4.634]**	[2.636]	[3.699]
Polity2	0	0.168	0.168	0.168	0.168	0.237	0.237	0.157	0.157
	[0.000]	[0.064]***	[0.088]*	[0.064]***	[0.088]*	[0.135]*	[0.143]	[0.072]**	[0.103]
Log Land Area	0.244	1.61	1.61						
	[0.317]	[0.368]***	[0.497]***						
Clustered Errors			Yes		Yes		Yes		Yes
Observations	300	308	308	308	308	107	107	201	201
R-squared	0.43	0.33	0.33	0.33	0.33	0.35	0.35	0.39	0.39

Weighted Least Squares regressions, weight = (number of observations per country-type)$^{-1}$. Robust standard errors in brackets. Constants are omitted. Clustered Errors are at country and quality level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10
Inequality and Population: Dynamic Panel with Country-Specific Effects

Dependent variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Population	-4.500**	-0.823	-3.990*	-3.576***	-5.744***	-0.58	-5.289***	-1.923
	[2.26]	[1.12]	[2.32]	[1.25]	[1.26]	[0.97]	[2.05]	[1.63]
Log Density		-3.644**				-2.738*		-3.058*
		[1.75]				[1.62]		[1.74]
Consumption Dummy	-2.475	-2.322	-2.88	-3.291**	-5.844***	-0.58	-4.302	-4.458
	[3.58]	[3.25]	[3.54]	[1.67]	[1.78]	[2.10]	[3.66]	[3.84]
Gross Dummy	3.987	4.111	3.228	4.546***	2.07	5.411***	-84.53***	-77.06***
	[3.29]	[3.44]	[3.06]	[1.31]	[1.73]	[1.95]	[31.5]	[28.2]
Full Democracy Dummy						-28.37		
						[22.8]		
Log Population * Full Democracy						0.271		
						[1.31]		
Log Density * Full Democracy						3.629**		
						[1.72]		
Log Population * Gross							5.198***	4.743***
							[1.88]	[1.64]
Log GDP per capita	-3.141	-1.826	-1.658	-4.374***	-5.707***	-0.34	-3.761*	-4.247**
	[3.61]	[2.71]	[2.99]	[1.31]	[1.34]	[1.58]	[2.22]	[1.94]
Openness	-0.0770*	-0.0729*	-0.067	0.0146	0.014	-0.0579**	-0.0464	-0.04
	[0.040]	[0.042]	[0.048]	[0.018]	[0.021]	[0.025]	[0.031]	[0.038]
ELF in 1985	9.588*	10.01**	9.591*	0.912	-4.333	10.19***	7.583	7.21
	[5.66]	[5.01]	[5.33]	[2.94]	[3.41]	[3.24]	[4.94]	[5.31]
Democracy Index	0.422**	0.324*	0.366	0.372***	0.441***	0.315*	0.368**	0.413**
	[0.19]	[0.19]	[0.23]	[0.14]	[0.11]	[0.17]	[0.18]	[0.18]
Log Land Area	3.501**		3.228*	2.907***	5.221***		2.955*	
	[1.62]		[1.75]	[1.09]	[1.03]		[1.61]	
Regional Dummies				Yes				
Legal Origine Dummies					Yes			
Observations	498	498	498	498	498	968	498	498
Number of country_type	119	119	119	119	119	168	119	119

All columns use Blundell-Bond GMM level equation dynamic panel method with optimal two-stage estimation. Log Population, Log Density, Log GDP per capita and interaction variables are treated as predetermined, except in Column (4) where they are treated as endogenously determined. Instruments include lags 1 to 5. The sample covers countries with democracy index less than 10, except in Column (7). Standard errors in brackets are clustered by country and corrected using Windmeijer's method. Intercept is omitted.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 11
Inequality and Population Concentration in Non-Democratic Countries
(Polity2<10)

Dependent Variable: Gini	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS				IV			
	Full Sample		Gross	Net	Full Sample		Gross	Net
Log PopCon	-12.28**	-9.543*	-4.126	-11.72	-16.52**	-9.819	-8.344	-15.75
	[6.08]	[5.50]	[5.31]	[8.83]	[8.22]	[6.37]	[10.0]	[11.7]
Log Population	-1.874*	-0.339	0.881	-1.627	-2.376*	-0.374	0.374	-2.092
	[0.97]	[1.06]	[1.08]	[1.43]	[1.23]	[1.20]	[1.48]	[1.80]
Gross dummy	-0.998	-0.518			-1.063	-0.527		
	[1.79]	[1.64]			[1.78]	[1.64]		
Consumption dummy	-7.624***	-6.058***		-6.167***	-7.458***	-6.058***		-6.152***
	[1.76]	[1.71]		[1.77]	[1.77]	[1.71]		[1.78]
Log GDP per capita	-2.198	-2.594	0.479	-6.165***	-2.354	-2.608	0.0826	-6.245***
	[1.47]	[1.70]	[2.13]	[1.38]	[1.43]	[1.71]	[2.05]	[1.42]
Openness	0.00693	0.0126	0.00846	0.0341	0.00575	0.0125	0.0075	0.0312
	[0.022]	[0.022]	[0.018]	[0.034]	[0.022]	[0.023]	[0.017]	[0.034]
ELF in 1985	5.338	3.031	14.84**	-1.213	4.267	2.988	12.63*	-1.169
	[4.21]	[4.20]	[6.47]	[4.00]	[4.16]	[4.10]	[7.22]	[4.08]
Polity 2	0.197	0.213	0.265	-0.0435	0.218	0.214	0.318	-0.061
	[0.15]	[0.16]	[0.24]	[0.19]	[0.15]	[0.15]	[0.27]	[0.20]
Legal Origin dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies		Yes		Yes		Yes		Yes
China/India dummies		Yes		Yes		Yes		Yes
Countries	72	72	44	59	72	72	44	59
R-squared	0.61	0.69	0.82	0.76	0.61	0.69	0.82	0.76

Weighted Least Square regressions, weight = 1/(number of observations per country-type). Robust standard errors in brackets.
Intercepts are omitted. All errors are clustered at country level.

*** p<0.01, ** p<0.05, * p<0.1

Table A1
Descriptive Statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
Gini	1352	35.10	9.75	15.55	63.30
Log Population	1352	16.67	1.57	12.81	20.98
Consumption dummy	1352	0.11	0.31	0	1
Gross dummy	1352	0.43	0.49	0	1
ELF	1117	0.38	0.25	0.00	0.92
Log GDP per capita	1134	9.12	0.94	6.32	10.98
Polity 2	1255	6.52	5.65	-9	10
Log Land Area	1252	12.57	1.86	7.61	16.64
Log Density	1252	4.16	1.27	0.43	6.84
Openness	1272	70.34	41.23	7.94	289.53
North America	1352	0.08	0.27	0	1
Western Europe	1352	0.26	0.44	0	1
Latin America and the Caribbean	1352	0.11	0.31	0	1
Sub-Saharan Africa	1352	0.02	0.15	0	1
Middle East and North Africa	1352	0.01	0.09	0	1
East Asia and The Pacific	1352	0.11	0.31	0	1
South Asia	1352	0.05	0.21	0	1
Eastern Europe and Central Asia	1352	0.36	0.48	0	1

Figure 1

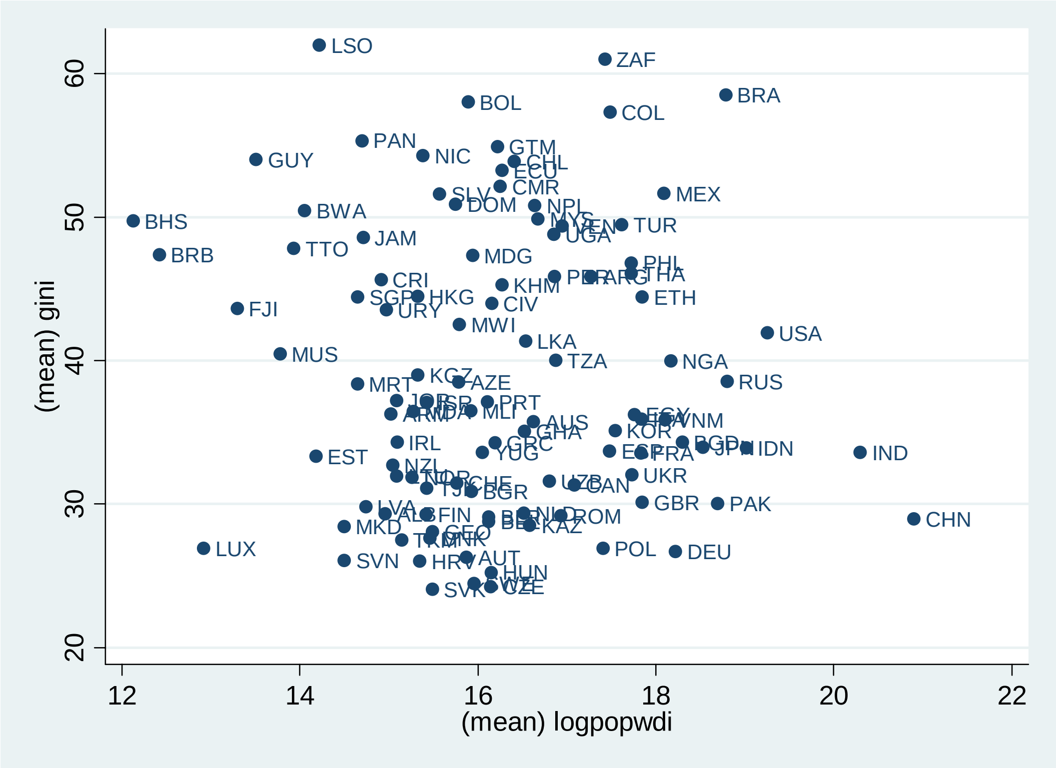

(mean) gini
60
50
40
30
20
12
14
16
18
20
22
(mean) logpopwdi
LSO
ZAF
BRA
BOL
COL
PAN
GTM
NIC
GUY
CHL
ECU
CMR
SLV
MEX
DOM
NPL
BWA
BHS
MYS
VEN
TUR
UGA
JAM
TTO
BRB
MDG
PHL
PER
ARG
THA
CRI
KHM
SGP
HKG
ETH
CIV
FJI
URY
MWI
USA
LKA
MUS
TZA
NGA
KGZ
AZE
RUS
MRT
ISR
PRT
ARM
MDA
MLI
EGY
VNM
AUS
GHA
KOR
IRL
GRC
BGD
JPN
IDN
IND
EST
YUG
ESP
FRA
NZL
NOR
UKR
CHE
TJK
UZB
CAN
BGR
GBR
PAK
LVA
ALB
FIN
BEL
NLD
ROM
MKD
KAZ
CHN
GEO
TKM
DNK
LUX
POL
DEU
AUT
SVN
HRV
HUN
SVK
SWE
CZE

Figure 2: Timeline

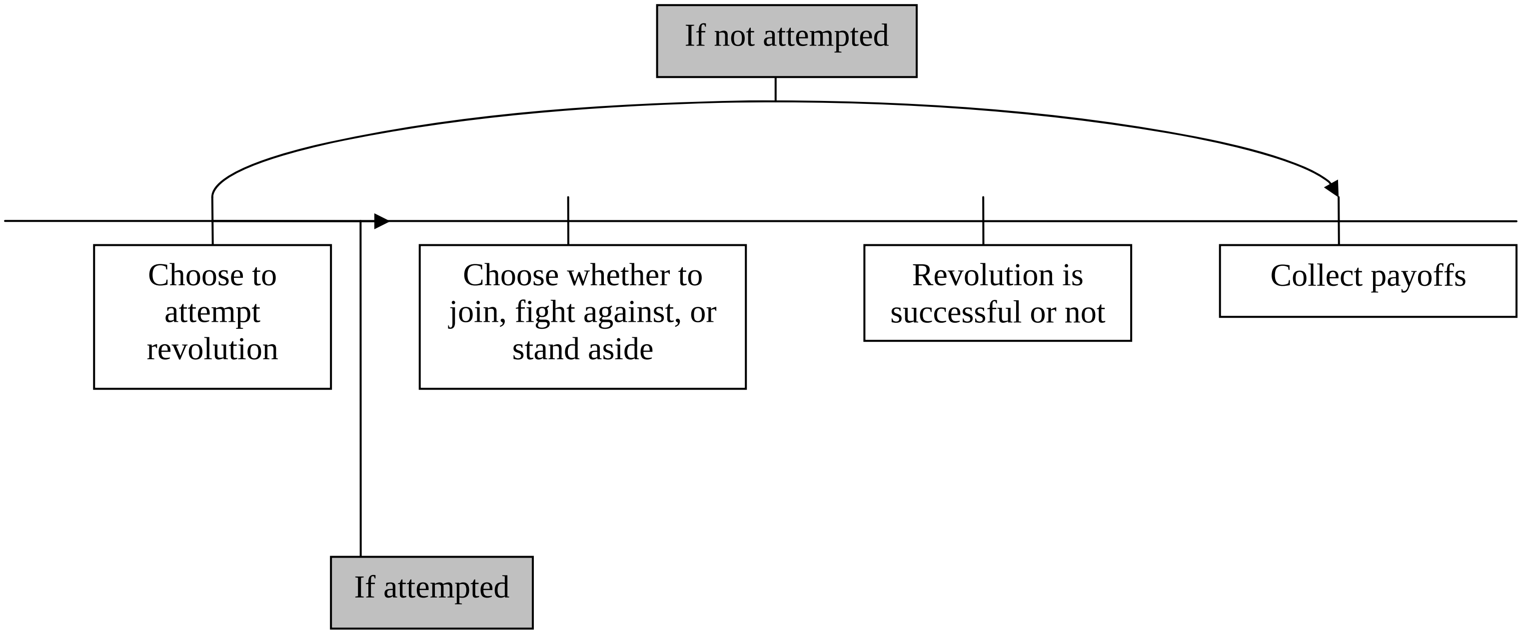


Figure 3

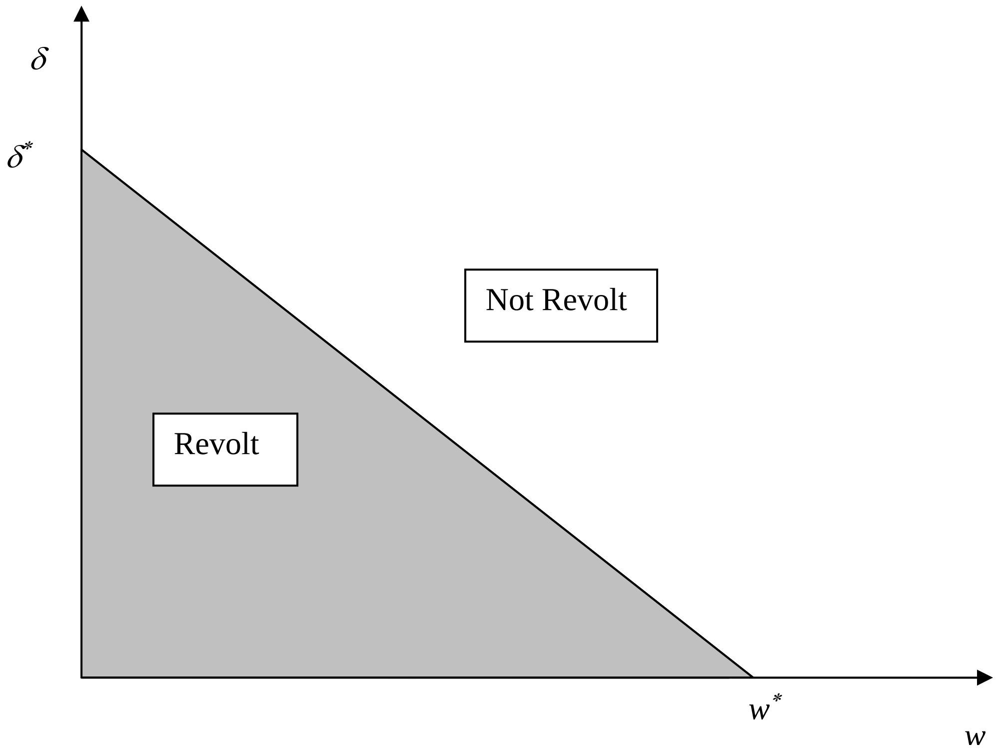
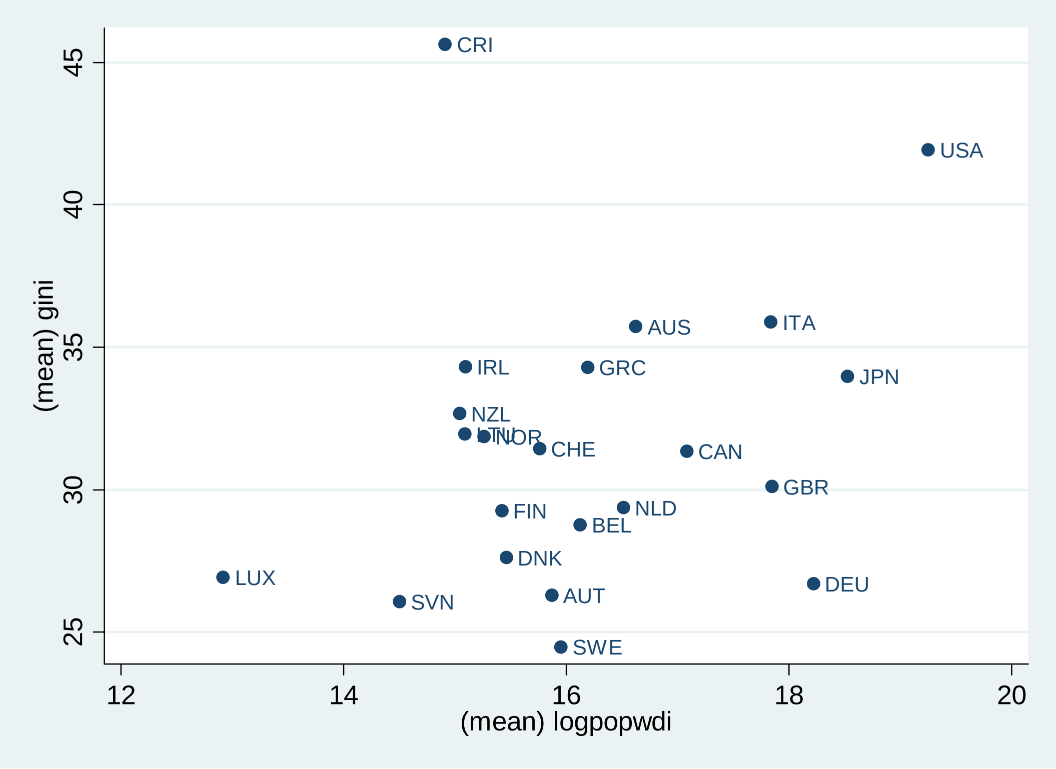


Figure 4



(mean) gini
45
40
35
30
25
12
14
16
18
20
(mean) logpopwdi
CRI
USA
AUS
ITA
IRL
GRC
JPN
NZL
LTU
NOR
CHE
CAN
GBR
FIN
NLD
BEL
DNK
LUX
DEU
SVN
AUT
SWE

Figure 5

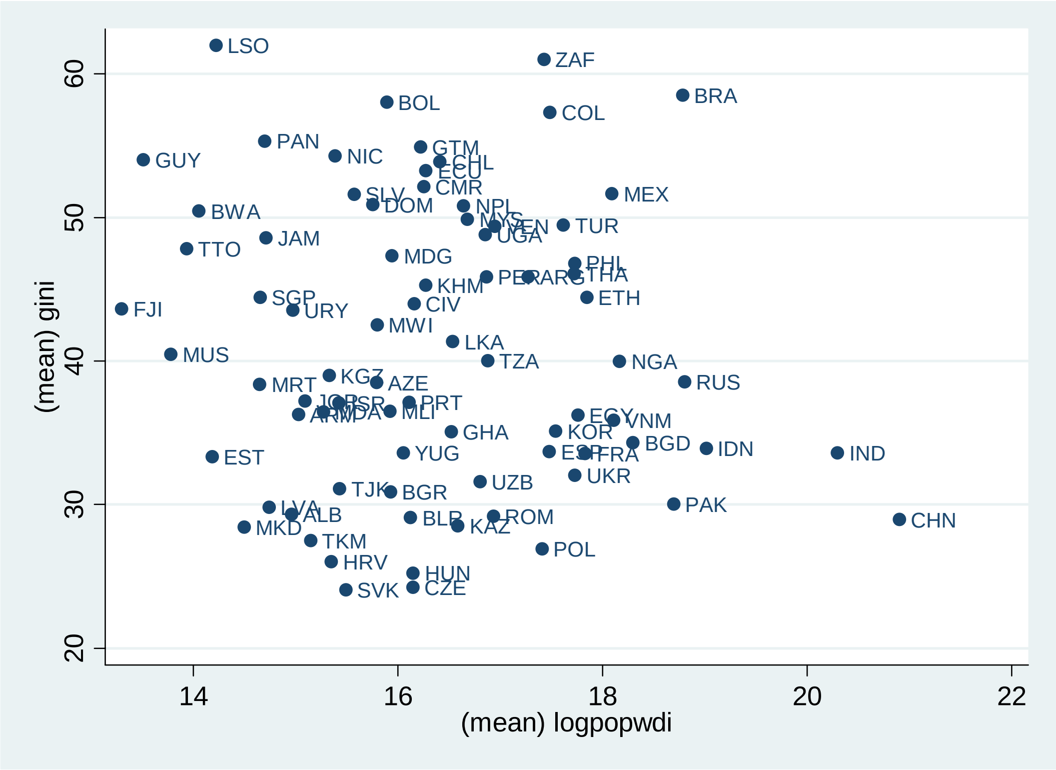
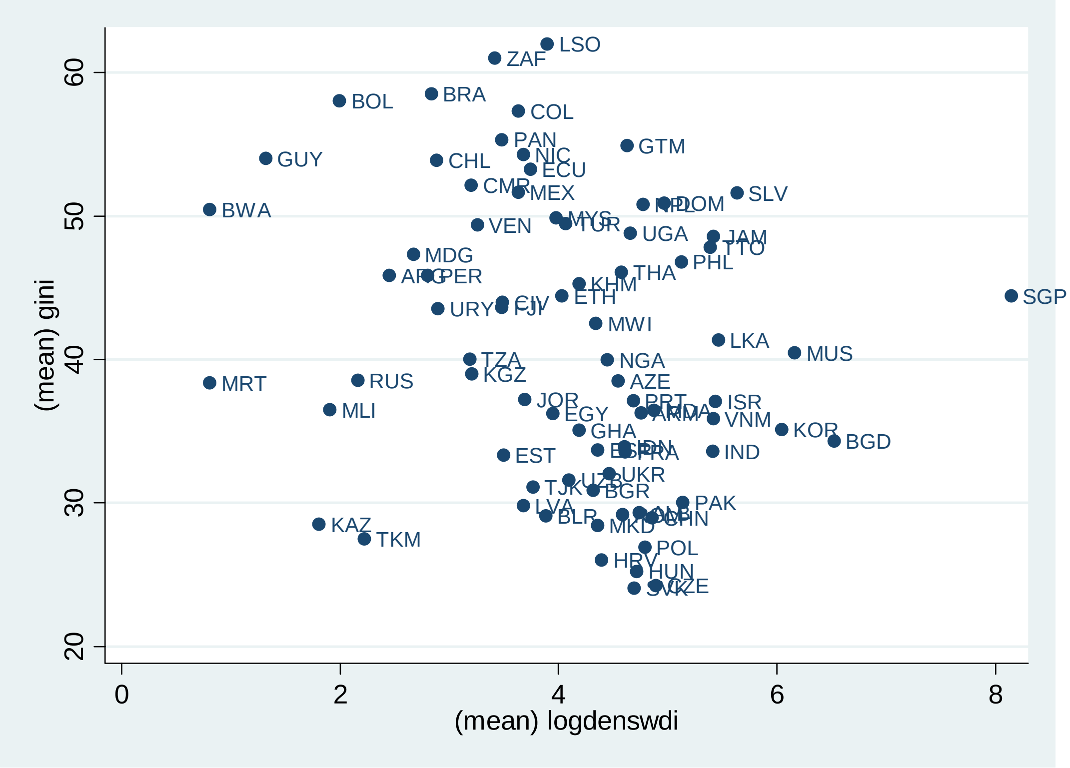

60
50
40
30
20
(mean) gini
14
16
18
20
22
(mean) logpopwdi
LSO
ZAF
BRA
BOL
COL
PAN
GUY
NIC
GTM
CHL
ECU
CMR
MEX
SLV
DOM
BWA
NPL
MYS
VEN
TUR
UGA
JAM
TTO
MDG
PHL
PER
ARG
THA
KHM
ETH
SGP
FJI
URY
CIV
MWI
LKA
MUS
TZA
NGA
KGZ
AZE
RUS
MRT
JOR
ISR
PRT
ARM
MDA
MLI
EGY
VNM
GHA
KOR
BGD
IDN
IND
EST
YUG
ESP
FRA
UKR
TJK
BGR
UZB
PAK
LVA
ALB
BLR
ROM
CHN
MKD
KAZ
TKM
POL
HRV
HUN
CZE
SVK

Figure 6


(mean) gini
(mean) logdenswdi
20
30
40
50
60
0
2
4
6
8
LSO
ZAF
BOL
BRA
COL
PAN
GTM
GUY
CHL
NIC
ECU
CMR
MEX
SLV
BWA
NPL
DOM
MYS
TUR
VEN
UGA
JAM
TTO
MDG
PHL
THA
ARG
PER
KHM
ETH
SGP
CIV
FJI
URY
MWI
LKA
MUS
TZA
NGA
KGZ
MRT
RUS
AZE
JOR
PRT
ISR
MDA
ARM
EGY
VNM
MLI
KOR
GHA
BGD
ESP
IDN
FRA
IND
EST
UKR
UZB
TJK
BGR
PAK
LVA
ALB
CHN
KAZ
BLR
MKD
TKM
POL
HRV
HUN
CZE
SVK