Vote or Fight?*

David K. Levine[†], Cesar Martinelli[‡], Nicole Stoelinga[§]

first version: June 23, 2021

Abstract

When political power is indivisible, voting is a substitute for fighting. However, voting is only meaningful if the loser chooses not to fight. We study a theoretical model of this substitution, and assess empirically whether it is true that the same economic fundamentals that determine fighting determine voting as well. They do. We introduce a number of theoretical and empirical innovations. We use a recently developed method of analyzing conflict resolution functions to develop robust theoretical results. We introduce a new measure, income relative to the global frontier, and explain why it matters theoretically and empirically. We also establish the stylized fact that fighting (and voting) depend on income relative to the global frontier, but not on how high that frontier happens to be. An important take-away is that reducing global inequality through economic development is essential to reducing conflict and increasing democracy.

Keywords: game theory, democracy, civil war, conflict, voting

*We would like to thank Rohan Dutta, Salvatore Modica and most especially Andrea Mattozzi. We are grateful to MIUR PRIN 20103S5RN3 and the Leverhulme Trust for financial support.

[†]RHUL Economics and Emeritus WUSTL; david@dklevine.com
[‡]Economic Department, George Mason University; cmarti33@gmu.edu
[§]Max Planck Institute for Research on Collective Goods; stoelinga@coll.mpg.de

1. Introduction

We study the problem of two parties contending for political power. Our basic premise is that political power is difficult to divide: that one party must win. Ultimately power must go to the stronger party, but there are two distinct ways in which strength can be measured. One is peacefully through voting and the other is violently through fighting. However elections are only meaningful if the loser accepts the results and does not choose to fight: hence our point of departure that voting and fighting are substitutes. The goal of this paper is to develop a simple model of this substitution and use it to assess the probability that there is fighting - by which we mean the attempt to seize power through force - versus voting - by which we mean elections in which the loser will respect the result.

Our point of departure is a theoretical model of voting and fighting as substitutes. A key conclusion of the model is that the same economic fundamentals that explain the prevalence of fighting also explain the prevalence of democracy. We examine whether this is true empirically and find that it has strong support in the data.

Our model of voting and fighting has two innovative aspects. First, the study of conflict has been handicapped by the fact that the marginal increase in the probability of winning tends to be larger when the two parties are more evenly matched and the resulting equilibrium is in mixed strategies. These equilibria are difficult to compute and analyze (see, for example, Ashworth and Bueno De Mesquita (2009)). Our results are driven by recent theoretical developments, most importantly by Ewerhart (2017), showing that, while the mixed strategy equilibria may be difficult to compute, they are unique, and the crucial characteristics of that equilibrium, which is to say the probability of winning and the equilibrium expected effort, do not depend on the details of the conflict resolution function. Second, we provide a theory compatible with the data, about how income determines the probability of both fighting and democracy.

Specifically, our key focus is on the role of income in determining voting versus fighting. This is known to be an important and robust explanatory variable for explaining civil war, and casual empiricism (as well as our data) indicate that it is an important explanatory variable for both voting and fighting. However, by examining data on civil conflict, income, and voting from 1815 to the present, we establish a key stylized fact that must be contended with: the impact of income over time is quite different than in the cross-section. Over time, the enormous increase in income has not resulted in a greatly reduced probability of fighting. This is shown clearly below in Figure 1 where if anything the probability of fighting has increased with income.

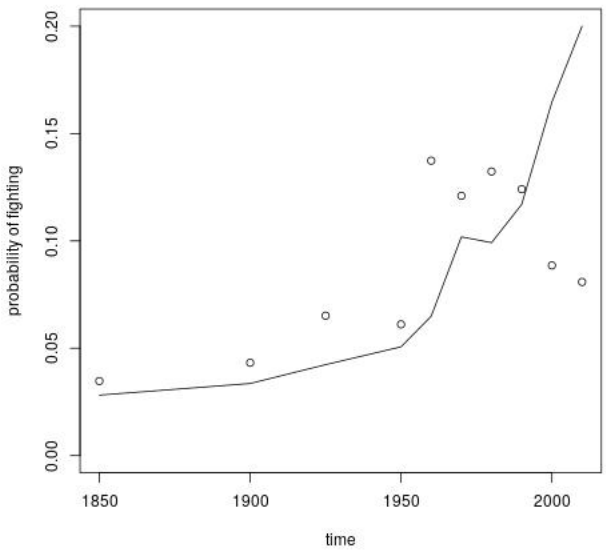

Figure 1: Fighting versus time

Dots: probability of fighting; curve average per capita GDP normalized to be equal to 0.20 in 2010.

By contrast, in the cross section, if we take the ratio of per capita GDP to the frontier country with the highest per capita GDP, we find that the probability of fighting is a strongly decreasing.

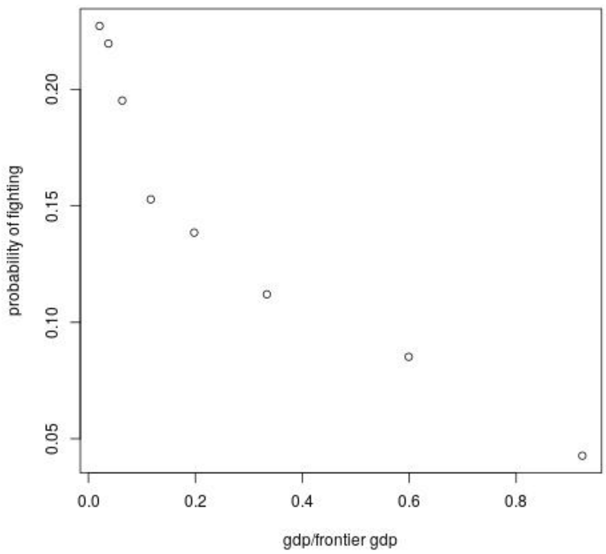

Figure 2: Fighting versus Economic Efficiency

In our model the decision of whether to fight or vote is driven by a benefit cost analysis. Both voting and fighting are costly - either turning out voters or turning out soldiers. However fighting also creates battle damage: by our estimate $100 spent on soldiers creates around $2000 of damage. A key innovative feature of our theoretical model is an explanation of why this benefit cost analysis should depend on relative per capita GDP. Our hypothesis is that countries with low per capita GDP operate the frontier technology inefficiently and that battle damage increases this inefficiency. That is, a poor country looks like a rich country with battle damage. As inefficiency increases, the additional harm done by battle damage is diminished: a very inefficient country does not become that much worse off by fighting.

The benefit cost analysis is different for an incumbent than the opposition: the incumbent has a substantial advantage and is consequently more willing to fight than the opposition. This means that there are three possibilities: if the opposition is willing to fight then neither party is willing to abide by the result of the election and there will be a fight. If the incumbent is not willing to fight neither party is willing to fight and there will be voting. Finally, if the incumbent is willing to fight and the opposition not then the incumbent will remain in power without a meaningful vote: there will be a dominant party.

With these ingredients we assume that countries differ only in their economic efficiency. We then assess the impact of inefficiency on fighting. Using the parameters estimated from that analysis we ask whether those parameters also explain the incidence of voting. They do as shown below in Figure 3.

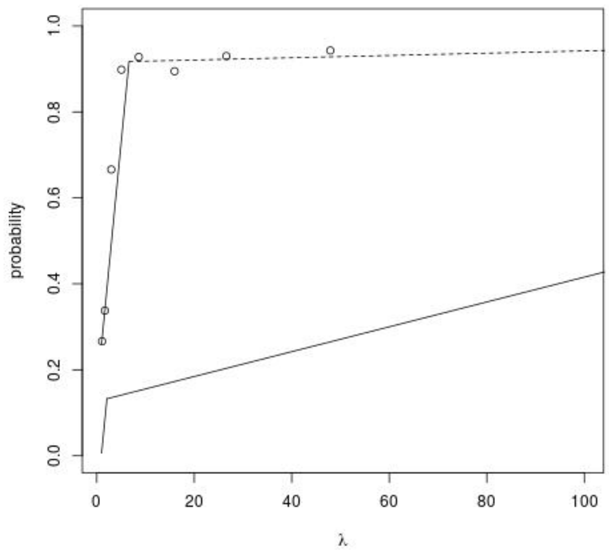

Figure 3: Fighting (lower line) and Absence of Voting (upper line)
λ is the lagged ratio of frontier per capita GDP to country per capita GDP

The lower curve is the estimated probability of fighting. The upper curve is one minus the probability of voting. The upper solid line is also estimated from the data on fighting: it is a pure out of sample forecast of voting based on conflict data. The upper dotted line is estimated from the voting data: there is no relevant fighting data to tell us what this might be. As can be seen relative per capita GDP (λ) is extremely important in explaining both fighting and voting and does so in a compatible way.

Having a theoretical model is important for assessing data. As an example, Acemoglu et al (2008) argue that empirically income does not lead to democracy. Generating data from our model we can use the Acemoglu et al (2008) procedure on our artificial data. Although we know that causality in this data runs from income to democracy, we find exactly the same empirical results that Acemoglu et al (2008) argue proves the opposite.

We also show that the size of the prize obtained by the stronger party is uncorrelated with the relative per capita GDP and consequently we can assess the impact of oil and polarization on the size of the prize. It increases the size of the prize in both cases. In the case of oil, the results are consistent quantitatively with the model. In the case of polarization, there is a discrepancy between the size of the prize estimated from the fighting and the voting data. This may be due to the fact that polarization has an impact on parameters other than the size of the prize.

There is an important message from this analysis. We take it that voting is good and fighting is bad. The conclusion is that to move the world to a state of greater voting and less (internal) fighting it is crucial to reduce inequality among nations. That is: the problem lies in nations that lie far from the frontier and helping them move towards the frontier would lead to great improvements in peacefulness along with the obvious economic benefit. However: they must move towards the frontier from economic development - increased productive efficiency -, which increases the cost of fighting, and not, for example, through subsidies or foreign aid, which may increase the size of the prize.

2. Literature Review

We discuss the literature on the relationship between income and both democracy and civil war below in section 10. Before reviewing any literature we want to emphasize that our empirical analysis is theory driven and why this makes a difference.

First, our definition of both fighting and voting differ from standard definitions of civil war and democracy. In particular fighting means attempting to overthrow the government by violence. This excludes regional civil wars but includes coups and coup attempts that do not involve widespread violence. Voting means an election in which the incumbent may lose and will allow the opposition to take power if it does. It does not depend on the extent of the franchise as in Acemoglu and Robinson (2001). In particular while, for example, the enfranchisement of women in the US and UK improved democracy immeasurably, it did not change the fact that power was already being passed peaceably from one party to another through elections.

Second, theory tell us about measurement. Probably most important is it tells us that it is per capita GDP relative to the frontier that matters not absolute per capita GDP. It tells us how long periods should be: the length of period should be the time a government is typically in power before facing reelection. (As a practical matter this is five years.) Finally, it tells us that during five year periods we should measure five year GDP as an average, not as a single year out of the five. The reason is that GDP matters because of the expectations it creates for the parties about the costs and benefits of conflict. Last year's unusually large or small GDP is scarcely reason to change these expectations.

Our work here contributes both to the literature on civil war and that on democracy. As indicated we discuss democracy below section 10. Fearon (1995) provides a classical analysis of the sources of civil war and it is useful to place our model into his framework.

- Bounded rationality: this is present in our model in the form of random optimism and pessimism about battle damage. In our model, fighting is driven by optimism and voting by pessimism.

- Agency problems: this is not part of our model.

- Asymmetric information: this is not part of our model. For a model that does incorporate asymmetric information, see Dal Bo and Powell (2009). This is a signaling game in which the government has private information about the size of the spoils. Like our model, this leads to a theory in which the probability of fighting is higher when income is low.
- Commitment problems: this is implicitly present in our model - it is essential that the parties not be able to commit to respecting the results of an election.
- Indivisibilities: this is our main assumption, that power cannot be shared. See, for example, Hirshleifer, Boldrin and Levine (2009) on why this is so.

Finally, there are several alternatives beyond income offered to explain the incidence of fighting. Esteban, Mayoral and Ray (2012) bring a theoretical model of polarization due to ethnicity to study data on ethnic conflict. There have been several studies concerning the relationship between trade and civil war: Besley and Persson (2008) focus on export and import prices, while Dal Bo and Dal Bo (2011) study a general equilibrium model. Those interested in this interesting and diverse literature will find the excellent survey article of Blattman and Miguel (2010) a helpful guide.

3. The Model

There are repeated games between two parties: in each period one party is the incumbent party i and the other is opposition party o. As a result of the game their roles may be the same or reversed in the next period. Specifically, both parties are myopic and in each period there is a game determining which party will take power. The winning party is the incumbent in the next period and wins a prize in the current period that both parties value equally.

First a random shock ϵ, the *degree of pessimism*, iid over time is drawn and is common knowledge. Then the opposition moves and may either *challenge* the incumbent or *concede*. If the opposition concedes the game ends and the incumbent party remains in power. (There may or may not be a non-credible "show" election: this does not matter to the game.) If the opposition challenges then the incumbent moves and may either hold a credible *election* or trigger a fighting. In other words, the challenge by the opposition is a credible commitment to fight if a credible election is not held. If a credible election is held it results in a winner and loser as described below. The loser then has a final opportunity to fight.

If there is an election the winner is determined stochastically depending on the *electoral effort* $e_k \geq 0$ provided by each party $k \in \{i, o\}$. If there is fighting any election results are discarded and the winner is determined stochastically depending on the *fighting effort* of the two parties $f_k \geq 0$.

For empirical purposes we will compare different countries at different times. We take the fundamental economic difference between countries to be per capita

GDP, which we denote by γ. This serves as a scaling factor for the benefits and costs of conflict. Specifically, both groups value the prize equally as $\gamma V > 0$ per capita where $0 \leq V \leq 1$. Effort is costly to both groups. The expected marginal cost of effort for group k is γB_k and γC_k for the election and fighting respectively. Note the assumption that in a higher income country the value of the prize is proportionately higher, but so is the opportunity cost of providing effort for a conflict. The realized expected direct cost to group k is $\gamma B_k e_k + \gamma C_k f_k$. In addition fighting creates costly battle damage: group k suffers an additional expected cost of $d(f_{-k})$ depending on the effort of the other group. We will provide more details about the damage function $d(f_{-k})$ below. We assume that the incumbent has an advantage in the contests in the sense that $B_i < B_o$ and $C_i < C_o$.

The outcome of each contest is determined by effort according to contest success functions. Specifically, we assume that the probability of k winning the contest does not depend on the scale of the conflict, only on the relative effort of the two parties. Hence, for $e_{-k}, f_{-k} > 0$ and $e_k \leq e_{-k}$ and $f_k \leq f_{-k}$, the probability that k wins is given by $P(e_k/e_{-k})$ for voting and $Q(f_k/f_{-k})$ for fighting respectively. As these function depend only on relative effort they should satisfy $P(1) = Q(1) = 1/2$. We assume that greater effort leads to greater success, so that these functions are weakly increasing and following Hirshleifer (1989) we assume that a small amount of additional effort is more likely to make a difference in a close contest than a one-sided one, so that these functions are weakly convex. Note that these assumption imply continuity on $[0, 1)$. More strongly, we requires that where the function is strictly increasing it is strictly convex. It will be convenient to abbreviate $P_0 = P(0), Q_0 = Q(0)$. If neither party provides effort each has an equal chance of winning.

Three examples of contest success functions $H(x)$ satisfying our assumptions are translations of the Tullock function $H_0 + (1 - 2H_0)\,(1/(1 + x^{-\alpha}))$ with $\alpha > 2$, translations of the serial contest success function $H_0 + (1 - 2H_0)(1/2)x^{\alpha}$ with $\alpha > 1$, and translations of the all-pay auction in which the probability of winning for $x < 1$ is H_0, where in each case $0 \leq H_0 \leq 1/2$.

We assume that $P_0 > 0$ meaning that if there is an election then there is some chance of success regardless of effort. As a practical matter we believe that there is: in particular, we believe that there is a difference between not contesting an election and providing no effort. If there are two candidates on the ballot, regardless of effort, unusual circumstances may intervene. For example, in January 1986 the Democratic presidential hopeful Gary Hart was polling nearly 46%.. His closest rival, Mario Cuomo pulled out of the race and he retained a commanding lead over his Democratic rivals until in May 1987 when photographs of himself with scantily clad women who were not his wife appeared in a number of newspapers, and he withdrew from the race. As a result Michael Dukakis became by default the Democratic candidate, and the Republican nominee, George H.W. Bush won in a landslide. Although in 1983 Edwin Edwards said "The only way I can lose this election is if I'm caught in bed with either a dead girl or a live boy" stranger things have happened.

Our notion of equilibrium is subgame perfection.

Productivity and The Damage Function

Different countries at different times have different levels of productivity. A key component of the model is how opponent effort creates damage in the form of lost output. We adopt a simple model of productivity differences. We imagine that at any moment of time there is an overall technology parameter representing the frontier economy: denote this by $g > 0$. Economies are not equally efficient however, and we imagine that an economy is characterized by how much time it takes to produce the per capita output g.. Specifically we denote this by $\lambda \geq 1$ where $\lambda = 1$ are countries at the technology frontier, and higher values of λ represent less efficiency in production - due to misallocation, monopoly, corruption, protectionism and other production inefficiencies. Hence $\gamma = g/\lambda$. Our model of battle damage is one in which opponent effort increases proportionally the length of time it takes to produce output. Specifically, we assume that the time to produce γ when there is battle damage is $\lambda + Df_{-k}$. Hence actual per capita output is $g/(\lambda + Df_{-k})$ and expected battle damage is given as

$$d(f_{-k}) = \epsilon\left(\gamma - \frac{g}{\lambda + Df_{-k}}\right) = \epsilon\gamma\frac{1}{\lambda/(Df_{-k}) + 1}.$$

where ϵ is the non-negative common random shock. The shock is assumed to have median equal to one meaning that if the true gain from fighting is zero for a party then there is a $50 - 50$ chance they prefer to fight.

As indicated the realization of the shock is common knowledge at the beginning of the period. It reflects the fact that battle damage is highly random. Indeed wars are often inflicted with random catastrophes, for example, involving the weather. Japan was saved from the overwhelming force of Genghis Khan in 1281 when a divine wind swept away Khan's navy. In 1941, despite complete surprise, poor leadership, complete lack of preparation, and the nervous breakdown of their supreme commander, the Soviet Union was saved from Hitler by the coldest winter in the 20th Century. As a result of this uncertainty and since fighting occur infrequently parties may be optimistic or pessimistic about how great the damage will be. We model this with the shock ϵ representing the degree of pessimism.

The critical feature of the battle damage function is that it is concave, equal to zero at $f_{-k} = 0$ and approaching γ as $f_{-k} \to \infty$. This means that a poorer country at a moment of time, corresponding to a larger λ, has lower marginal battle damage loss from effort. Hence, all other things equal, in the cross-section poorer countries find fighting less costly. By contrast overall economic progress as measured by g impacts all countries the same way, so that increasing income over time will not imply a secular decrease in the propensity for fighting.

4. Main Result

Define the incumbency advantages $\rho_e = B_o/B_i, \rho_f = C_o/C_i$ which are greater than one, the effort cost to damage ratio $r_d = C_0/D$, and the function

$$G(\rho) \equiv \left(\rho^{-1}Q_0 + (1-\rho^{-1})(1-Q_0)\right)\left(\lambda 2 r_d \rho/(1-2Q_0) + V\right).$$

It will later be shown that $G(1)$ is the benefit to cost ratio for rebelling for the opposition and $G(\rho_f)$ is that for the incumbent.

Theorem 1. *There are three cases:*

(fighting) $G(1) > \epsilon$*: there a fight in which there is probability* $\Pi_0 \equiv (1-\rho_f^{-1})Q_0 + \rho_f^{-1}(1/2)$ *that the opposition seizes power and the expected cost of effort (relative to GDP) to each of the two parties is the same and equal to* $\Pi_0 V$.

(dominant party) $G(1) < \epsilon < G(\rho_f)$*: the initial incumbent remains in power.*

(voting) $G(\rho_f) < \epsilon$*: there is an election in which there is probability* $(1-\rho_e^{-1})P_0 + \rho_e^{-1}(1/2)$ *that the opposition wins the election.*

Comparative statics are given by $dG/d\rho > 0$, $dG/d(\lambda r_d) > 0$, $dG/dV > 0$ *and for* $\rho < 2$ *also* $dG/dQ_0 > 0$.

What Determines Fighting and Voting?

Before turning to the proof of Theorem 1 we examine what it has to say. First it says that whether power is determined by fighting, voting, or belongs to a single dominant party depends (stochastically) only on the four fundamentals of fighting: the productivity adjusted effort cost to damage cost ratio λr_d, the value of the prize V (relative to GDP), the incumbent advantage in fighting ρ_f, and the degree of randomness in fighting Q_o. While the fundamentals of elections matter to who wins the election under voting they do not matter in the determination of institutions.

The basic comparative static for fighting and voting from Theorem 1 can be seen from studying the two cutoffs. The first cutoff $G(1)$ measures how attractive fighting is for the opposition and the second $G(\rho_f)$ which measures how attractive fighting is for the incumbent: $dG/d\rho > 0$ implies that fighting is always more attractive to the incumbent than the opposition. When the ratio λr_d is large both cutoffs increase reducing prospects for voting and increasing those of fighting. In particular, all other things equal, countries further behind the technology frontier are both less likely to be democratic and more likely to have fighting. A higher value of the prize V similarly increases both cutoffs reducing prospects of voting and increasing those of fighting. Increases in the incumbent advantage have no effect on the threshold between voting and a dominant party, but increases the cutoff for fighting meaning greater prospect of fighting. Finally increases the degree of randomness Q_0 unambiguously raises $G(1)$ increasing prospects of fighting. If the incumbent advantage is not too great, it also reduces the prospects of voting.

On Fighting

Theorem 1 also relates the parameters determining institutions to the nature of fighting if it occurs. The probability that the opposition wins is $\Pi_0 \equiv$

$(1-\rho_f^{-1})Q_0 + \rho_f^{-1}(1/2)$, which is increasing in Q_0 and, since $Q_0 < 1/2$, decreasing in ρ_f. That is, greater incumbent advantage reduces the chances of the opposition winning. Notice, however, that increasing ρ_f (holding fixed r_d) has no effect on the chance that fighting occurs. For example, more repressive state by increasing incumbent advantage reduces the chance that the opposition succeeds in fighting, but does not reduce the chance of fighting unless it also reduces r_d, for example by increasing the amount of damage inflicted on the opposition per given unit of force.

The second result says that the intensity of fighting as measured by the expected cost of effort (relative to GDP) is proportional to the probability of the opposition winning, and the factor of proportionality is exactly the size of the prize (relative to GDP).

5. Proof of the Main Theorem

As players are myopic it suffices to analyze the stage game. The proof of the main theorem then follows from a basic result on contests and some calculations. The basic result on contests is this:

Theorem 2. *Consider a two-party contest with prize V which is won by k with probability given by the increasing convex function strictly convex when strictly increasing $H(g_k/g_{-k})$ with $H(1) = 1/2$, where g_k is effort and the cost of effort is A_k. Define the disadvantaged party d as having $A_d \geq A_{-d}$, and define*

$$\rho = \frac{A_d}{A_{-d}}.$$

Then the disadvantaged party gets utility $u_d = H_0 V$ and wins with probability $\pi_d = H_0 + (1-2H_0)(\rho^{-1}/2)$ and the advantaged party gets

$$u_{-d} = \left(1 - H_0 - (1-2H_0)\rho^{-1}\right) V.$$

Expected efforts $\overline{G}_k$ are computed from $u_k = \pi_k V - A_k \overline{G}_k$ and in particular the expected cost of effort $A_k \overline{G}_k = (1-2H_0)(\rho^{-1}/2)V$ is the same for both parties.

This result is suprising and is neither obvious and nor easy to prove. A slightly weaker version was first shown by Ewerhart (2017) and all of the proofs are based on a crucial idea from Alcalde and Dahm (2007). This particular version follows from Theorem 11.7.1 in Levine, Mattozzi and Modica (2022) based on a similar result in Levine and Mattozzi (2021). We give a brief indication of why it is true. The first idea is that an equilibrium exists. Given this, the second idea is that at the bottom of the support of an opponent's strategy a party faces the expected value of convex utility functions: this must be convex and that means that a party cannot be optimizing at the bottom of the support of the opponent unless they have they same bottom or the bottom is zero. The same support can be ruled out, and the idea extends to show that both must have the bottom at zero. This "bidding down to zero" due to convexity of the contest success function captures Hirshleifer (1989)'s intuition about contests and is reminiscent of the ideas in the derivation of equilibrium in the all-pay auction. Notice that both parties will not bid zero in equilibrium with positive probability, so this argument also establishes that the equilibrium must be in mixed strategies.

Bidding down to zero enables us to conclude that one of the parties k gets H_0V and that the equilibrium utility of the other depends only on the probability that k plays zero. The final idea is to use the method of Alcalde and Dahm (2007) to construct another game in which $-k$ has proportionally higher costs, and k instead of having an atom at zero plays bids with proportionately higher probability. This is an equilibrium of the modified game. Finally, following Ewerhart (2017), we show that this modified game must be symmetric and that this implies that both players must get the same utility, that is H_0V, and have equal probability one-half of winning. Mapping the equilibrium of the modified game back to the original then gives the desired winning probabilities and utilities.

This basic result applies immediately to the election contest. A crucial fact about the fighting contest is that battle damage has no effect on a party's incentives as it depends only on the actions of the other party. Hence the basic result applies also to fighting: we can compute equilibrium without battle damage, then subtracting the cost of battle damage from the expected utility of each party. Applying the basic result then yield the following values for the utility, expected effort and probability of winning.

Theorem 3. *Equilibria of the contests satisfy*

	expected effort	*probability of winning*
election incumbent	$(1-2P_0)(\rho_e^{-1}/2)V/B_i$	
election opposition	$(1-2P_0)(\rho_e^{-1}/2)V/B_o$	$P_0+(1-2P_o)(\rho_e^{-1}/2)$
fighting incumbent	$(1-2Q_0)(\rho_f^{-1}/2)V/C_i$	
fighting opposition	$(1-2Q_0)(\rho_f^{-1}/2)V/C_o$	$Q_0+(1-2Q_0)(\rho_f^{-1}/2)$

The main theorem now follows from the result for the stage game:

Theorem 4. *There are three cases:*

(fighting) $G(1) > \epsilon$*: elections do not matter and no effort is expended on them, there is fighting. Define*

$$\xi(\rho) = \gamma V\left[\left(1 - Q_0 - (1-2Q_0)\rho^{-1}\right) - \frac{\epsilon}{(\lambda r_d \rho)/(1/2 - Q_0) + V}\right].$$

Then the incumbent party gets $\gamma + \xi(\rho_f)$ *and the opposition party gets* $\gamma + \xi(1)$ *and wins with probability* $Q_0 + (1-2Q_0)(\rho_f^{-1}/2)$.

(dominant party) $G(1) < \epsilon < G(\rho_f)$*: the opposition party concedes and gets nothing while the incumbent gets* $\gamma(1+V)$*. If an election were to take place and the opposition were to win there would be fighting.*

(voting) $G(\rho_f) < \epsilon$: elections take place and the winner takes office. The incumbent party gets $\gamma\left(1 - P_0 - (1-2P_0)\rho_e^{-1})\right)V$, the opposition party gets $\gamma\,(1 + P_0V)$ and wins with probability $P_0 + (1-2P_o)(\rho_e^{-1}/2)$.

Proof. For ease of parsing expressions set $\eta = 1/2 - Q_0$. From Theorem 3 the expected gain from fighting for the opposition is

$$= \gamma V\left[Q_0 - \frac{\epsilon}{\eta^{-1}\lambda r_d + V}\right] = \xi(1)$$

and that of the incumbent is

$$= \gamma V\left[\left(1 - Q_0 - (1-2Q_0)\rho_f^{-1})\right) - \frac{\epsilon}{\eta^{-1}\lambda r_d \rho_f + V}\right] = \xi(\rho_f).$$

Setting $\xi(\rho) = 0$ and solving for ϵ gives the expression for $G(\rho)$.

When the fighting decision is made the effort expended in the election is a sunk cost. The losing party will fight, then, when the expected utility from the fighting is positive and will not when it is negative. The factor γ is irrelevant. Observe that G is increasing in ρ so if the opposition prefers fighting so does the incumbent and there is a fight and elections are pointless. If the incumbent does not prefer fighting we are in the case of voting, as nobody is willing to fight, the opposition does not concede, the election is credible and no effort is made to overturn the result. In the remaining case the incumbent prefers to fight and the opposition does not, so the opposition concedes rather than commit to a fight it does not want. This is the dominant party case. □

While we consider P_0 equal to zero uninteresting, for completeness we describe the equilibria.

Theorem 5. *If $P_0 = 0$ then in case (voting) there are additional equilibria*

(additional voting) $G(\rho_f) < \epsilon$: it is an equilibrium for the opposition to concede with any probability, in which case o gets 0 and i gets γV, and if the opposition does not concede, the outcome is as described in case (voting) above. Conceding with probability 1 Pareto dominates all other equilibria.

6. The Data

We assess the impact of income on voting and fighting in several stages. In the first stage we examine data on the conflict resolution function and estimate Q_0 and the incumbent advantage. In the second stage we assess the impact of income on fighting and in the third stage the impact of income on voting. In the first stage we use data on the outcome of civil wars. In the second stage we use data on income and the incidence of fighting and voting. In the third stage we use data on oil and natural gas value to assess the impact on the value of the prize.

The data spans from 1815 to the present. The theory tells us the appropriate length of a period: as elections are not held every year, a reasonable length of period is five years as this is about the time between major elections in places such as the US and UK. Hence we aggregate our data into half decade averages. The periods of World War I and World War II were excluded as the model does not apply to countries involved in heavy warfare with other countries. In all there are 4160 obervations on 215 countries. Summary statistics and sources for the variables discussed below are in Table 1.

Civil War Outcomes

The data on civil wars is from the Correlates of War (COW) project on civil conflicts as described in Dixon and Sarkees (2016). We use the intra-state wars dataset and focus on the on wars coded as a civil war over central control as this is the type of fighting contemplated in the model. We construct three variables: the ratio of the effort of the weaker side to the stronger, the ratio of the expected effort of the incumbent to that of the opposition, and a binary variable indicating which side won.

As a proxy for effort we started with the maximum in theater forces for that party during the conflict. If this data was missing for either party the observation was excluded. However, person power is not all equal, and some soldiers are more effective than others. In particular the incumbent will generally have better trained and equipped forces, better bases and so forth. Based on Dupuy (1986)'s estimate that in 1941 German troops were about three times as effective as Russian troops, we took incumbent effort f_i to be triple the number of in theater forces while for the opposition f_o we took it to be the number of in theater forces.

The outcome variable is binary being equal to one if the weaker side won. If both sides had equal strength this was scored as an 0.5 chance of the either side winning. If neither side won then the state remained in control so this is classed as a win for the state actor, that is the incumbent. To be consistent with our five year period length if the civil war lasted more than five years we broke it into five year periods with the state actor winning until the five year period in which the war ended.

Fighting and Income

Population and GDP data are from the Maddison Project as described in Bolt and van Zanden (2020) and the World Bank in 2011 US dollars.[1] The data on civil conflict from the COW project is augmented with data on coups from the Wikipedia list of coups and coup-attempts by country.[2]

The fighting state should include substantial unrest designed to or with the possibility to bring down the government (general strikes, widespread protests) as well as attemped or successful coup d'etats. A period is a fighting period if there is fighting at any point during the period. If there is no fighting, the

[1]There is a slight difference between the Maddison data and the World Bank data where they overlap. We corrected this by using the ratio of US values in 1960.

[2]See: https://en.wikipedia.org/wiki/List_of_coups_and_coup_attempts

period is classified according to whether or not there is a dominant party or there is voting as described below. A fight is said to occur in a five year period if there is a civil war or the continuation of a civil war lasting more than five years, a coup, or a coup attempt. In our data, there are 561 cases of fighting: 151 civil wars, 244 coups and 144 attempted coups. While the COW dataset records the duration of most conflicts, this data is not available for all coups; when missing the duration of a conflict event is set to one year.

Income determines GDP relative to the frontier $\tilde{\lambda}$. The frontier country is always taken to be the country with the highest per capita GDP and at least 0.5% of world population. This excludes small countries with a lot of mineral wealth (the Middle East) and countries with large banking sectors (Ireland, Luxemberg, and so forth). The frontier country is the UK until 1880 when it switches to the USA. This is consistent with the history of technology. Countries with per capita GDP higher than the frontier country are assumed to have $\tilde{\lambda} = 1$, that is, to be on the frontier.

Beliefs by the parties about λ determine their beliefs about battle damage. The natural measure is $\tilde{\lambda}$ from the previous period, but we must measure what $\tilde{\lambda}$ "would have been in the absence of fighting" and if there was heavy fighting in the previous period $\tilde{\lambda}$ may be artificially depressed. Our base assumption is that recovery from a fight is fast: that recovery from battle damage is swift - on the order of five years - is supported by data from World War II. For example,[3] the population of Hiroshima in 1940 was 1.9 million and in 1950 - five years after the nuclear bombing - it was 2.1 million. Similarly, despite the immense battle damage, the US civil war seems to have had little impact on subsequent economic success.

For this reason, if heavy fighting last period depressed $\tilde{\lambda}$, a better measure is the last value of $\tilde{\lambda}$ before a fight broke out. However, fighting does not depress $\tilde{\lambda}$ in all cases (for example a coup is unlikely to have much effect). Hence we take as our explanatory variable λ to be the smaller (more efficient) value of last period $\tilde{\lambda}$ and the last value of $\tilde{\lambda}$ before fighting broke out. For the sake of brevity we refer to this as the "lagged ratio of frontier per capita GDP to country per capita GDP".

Voting

Data on democracy is from the Varieties of Democracy (V-Dem) project as described in Pemstein et al (2024). Our basic measure of voting is based on their polyarchy index, which captures the extent to which electoral democracy is achieved in a country and is based on measures of freedom of association and expression; freedom of expresson, clean elections, suffrage and elected officials. This does not measure exactly what we want, as it places emphasis on the extent of the franchise which is not part of our theory. In the modern era, however, the franchise when there is one is generally universal and the index, which has

[3]http://www.demographia.com/db-japanpref.htm

been carefully constructed and well tested seems to do a good job in measuring voting.

To use the index, which ranges from zero to one, we need a cutoff indicating when voting takes place and when it does not. Most intermediate level cutoffs do a good job picking out countries that are clearly democracies such as the US and Western Europe and that are clearly not such as China and the Gulf States. Most intermediate cutoffs also do a good job for countries where there was a large change due to institutions such as Argentina, South Korea, Mexico and Taiwan. To use it, however, we need it to capture correctly poor countries that may hover on the edge of voting.

We did this by examining Benin as a case study. Benin is a poor country that has had a brief episode of democracy. The polyarchy index for Benin is shown below in Figure 4.

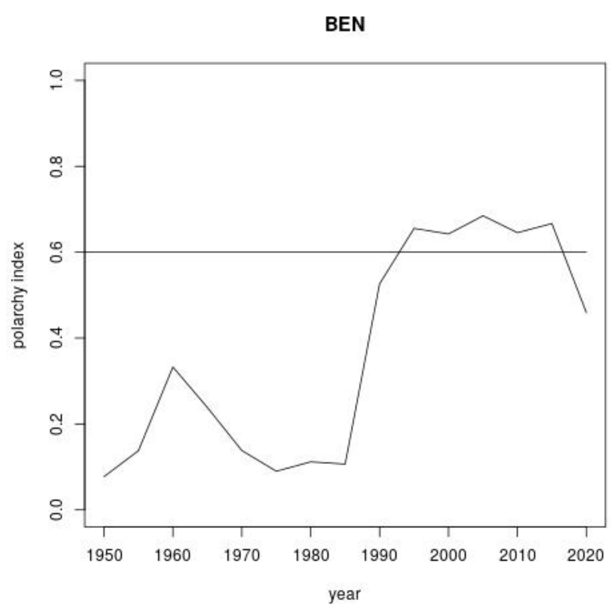

Figure 4: Polyarchy in Benin

From Wikipedia the salient facts about voting in Benin are these. Benin (then the Republic of Dahomey) achieved independence from France in 1960. There followed a period of civil strife culminating in the seizure of power by Lt. Col. Mathieu Kérékou who renamed the country as the People's Republic of Benin. Dominant party rule ensued until the inability to pay the army and a banking collapse resulted in an agreement for a new constitution and the further renaming of the country as the Republic of Benin. An election was held in 1991 and the incumbent Kérékou lost to Nicéphore Soglo who took power. There ensued a number of elections in which the winner was often not the incumbent and was able to take power. This lasted until Patrice Talon was elected in 2016. Talon then pushed through electoral reforms that disenfranchised the opposition and put the leading opposition leaders in prison.

From our perspective power in Benin was established by voting in roughly the period from 1991 to 2016: this is reflected in the fact that the polyarchy index in Figure 4 is considerably higher in those years than in other years. Taking a cutoff of 0.6, the horizontal line in figure 4 seems to capture the period of voting well and we take this as our basic cutoff.

As indicated there is a major issue with the polyarchy index before the modern era: the polyarchy index places substantial weight on the extent of the franchise and a measure of free fair elections. Our definition of voting, however, is that the loser should respect the outcome. While the franchise is important for democracy in the usual sense it is not so relevant for voting in our sense: the disenfranchisement of women is undemocratic but there is not a "men's" party and a "women's" party that take turns in power. Hence, while the polyarchy index seems to a good job of measuring voting in the cross-section it does poorly over time: the UK had peaceful transitions of power based on voting in the entire period and the US as well, except briefly around the time of the Civil War.

There are two methods of accounting for voting in the earlier period. One is to revise the index by reweighting the components. We experimented with this, but found that it is all too easy to get meaningless results. Instead of replacing a well thought out and well tested index with our own, we decided instead to adjust the cutoff for democracy in the earlier period. A simple and useful way of doing this is to use the polyarchy index in the frontier country as an indicator of the proper standard in earlier periods. Specifically, we chose the cutoff to be 95% of the polyarchy index in the frontier country or 0.6 whichever is smaller. This cutoff is shown below in Figure 5 below.

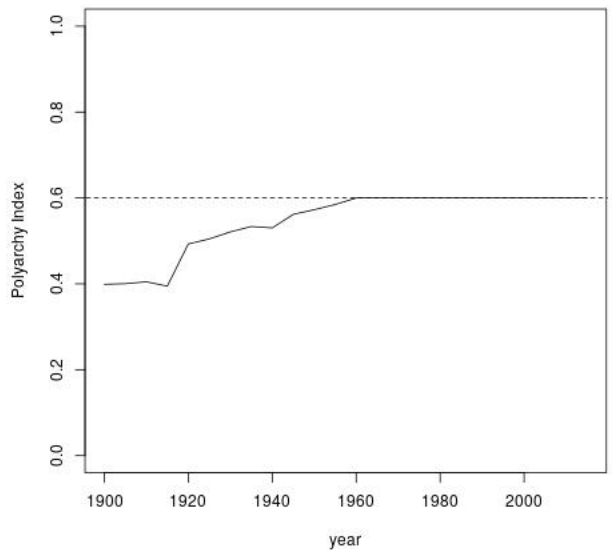

Figure 5: Polyarchy Index for the Frontier Country

Oil

The per capita value of oil and natural gas is from Ross and Mahdavi (2015) and extends through the five year period ending in 2009. We adjusted the year 2000 US dollars to 2011 US dollars using GDP deflator from the World Bank. Using the GDP per capita data is from the main data set.

Polarization

The polarization index is from Esteban, Mayoral and Ray (2012) and captures the distances between groups in a country based on similarity between languages and population share of a group and is available for 1960 to 2008.

Data Summary

The variables used in the study and summary statistics are in Table 1 below.

7. The Conflict Resolution Function and Q_0

We start by using data on the outcome of civil wars to assess the conflict resolution function. Recall that if k is the lesser effort of the two parties the conflict resolution function $Q(f_k/f_{-k})$ should depend upon relative effort and be weakly increasing and weakly convex. Moreover, as $f_k/f_{-k} \to 0$ it should converge to $Q_0 > 0$. Is this true and what is Q_0?

Variable	Mean	SD	Observations
Weaker to stronger f_k/f_{-k}	0.29	0.24	123
Incumbent to opposition Ef_i/Ef_o	3.96	N/A	123
Weaker side wins	0.28	0.45	123
Fighting	0.13	0.34	4160
No voting	0.74	0.44	3442
$\tilde{\lambda}$	8.95	10.91	3442
λ	8.63	10.24	3442
GDP/capita	9761	11514	1641
Oil & Gas/capita	918	4241	1641
Polarization	0.044	0.048	1368

Table 1: Descriptive statistics

To summarize the findings of this section, we find that the assumption of the probability of success depending only on relative effort and being weakly decreasing and weakly convex is consistent with the data. We estimate that Q_0, the chances of success against overwhelming odds, is about 7% and we estimate incumbent advantage to be about four to one.

The Conflict Resolution Function

In assessing the conflict resolution function since convexity is an issue and there are non-integer values of the endogenous variable we used a linear probability model. We studied a generalized conflict resolution function of the form $\tilde{Q}(f_k/f_{-k}, f_{-k})$ and approximated this by a quadratic. Below we report the regression results. None of the fitted values exceeded one and only three were negative being equal to -0.013. The results are in Table 2 below. Except for the constant term none of the coefficients are estimated with much precision.

	Quadratic		Force ratio only	
Variable	Estimate	SE	Estimate	SE
Constant	0.18	0.11	0.14	0.06
f_k/f_{-k}	0.94	0.61	0.49	0.16
$(f_k/f_{-k})^2$	-0.64	0.68		
f_{-k}	-1.02	0.76		
$(f_{-k})^2$	1.23	2.09		
$(f_k/f_{-k})f_{-k}$	-2.04	0.90		

Table 2: Conflict Resolution: Probability Weaker Party Wins

We are interested in whether the conflict resolution function depends only on the force ratio and is a convex function. As the cofficient on the quadratic term in the force ratio is negative we then reran the regression to see how well the model linear in the force ratio only works.[4] This time none of the fitted values are negative. Omitting the four other variables increases the sum of squared residuals by only 5.2%. Multiplying this by the sample size under the null hypothesis is drawn from an approximate chi-square with four degrees of freedom. The probability of getting 5.2% or larger is 16%, well above standard criteria for statistical significance. In summary: although the evidence is weak it is consistent with our theoretical assumptions.

The intercept term 0.14 is an estimate of Q_0, the chance of success against overwhelming odds.

Measurement Error

As we used a proxy for effort there is a problem of measurement error. The estimate of the constant term is negatively correlated with the estimate of the slope term, so measurement error in the force ratio will bias the constant term up leading to an overestimate of Q_0 (see, for example, Levine (1985)). To see if this is important we employed a robust technique for estimating Q_0 that is consistent in the face of measurement error.

The idea is to compute, for values of the proxy inverse force ratio $\varphi \equiv f_{-k}/f_k$ exceeding a threshold $\overline{\varphi}$, the percentage of the time the weaker party wins. In Online Appendix 1 we give conditions under which it is possible to choose $\overline{\varphi}$ as a function of the sample size so that asymptotically this converges in probability to Q_0.. The idea is that if we observed the true inverse force ratio $\tilde{\varphi}$ then the probability of lying above a threshold should asymptote to Q_0 so the same should be true for estimates based on the proxy force ratio.

The results of the estimation are graphed below in Figure 6 below.

[4]Note that linearity in the force ratio is not strictly convex, but if we cannot reject linearity we cannot either reject a slight among of strict convexity.

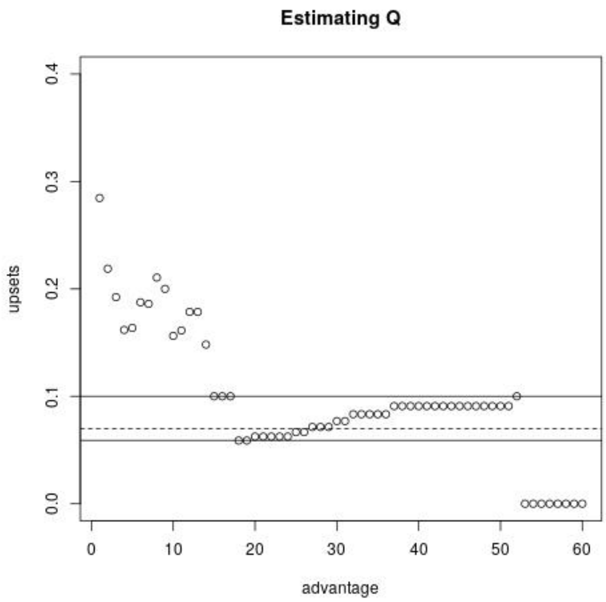

Figure 6: Force Ratios and Upsets

Naturally if we take $\overline{\varphi}$ too large there are too few observations to get a good estimate of Q_0. In Figure 6 it can be seen this happens at $\overline{\varphi} = 53$ where the number of observations has fallen from 10 to 5. For lower force ratios as the advantage increases the probability of the weaker party winning stabilizes in the range from 6% to 10% indicated by the solid lines in Figure 6. This is less than the linear probability model estimate of 14%. As there is more data in the lower range for force ratios around 30 we take 7% as a plausible value of Q_0, the dashed line in Figure 6.

Incumbent Advantage

Finally, using force data we can estimate incumbent advantage. This is given by $\rho_f = C_o/C_i$. From Theorem 3 equilibrium expected effort is given by $Ef_k = (1 - 2Q_0)(\rho_f^{-1}/2)V/C_k$ from which

$$\rho_f = \frac{C_o}{C_i} = \frac{(1-2Q_0)(\rho_f^{-1}/2)V/C_i}{(1-2Q_0)(\rho_f^{-1}/2)V/C_o} = \frac{Ef_i}{Ef_o}.$$

From the data on civil wars we can compute for each observation the force measured in numbers per capita. Averaging over the sample we find for the incumbent $Ef_i = 1.21\%$ and for the opposition $Ef_o = 0.31\%$ giving the estimate $\rho_f = 3.96$.

8. GDP and Fighting

Our estimation strategy is to assume that the only difference between countries lies in their time to produce λ. We have already used data on who won civil wars to estimate the chance of winning against overwhelming odds Q_0 and the data on force ratios to estimate ρ_f. We now use data on GDP per capita and the incidence of fightings to estimate the distribution of the shock ϵ and r_d and V. In the next section we ask whether these estimates can also explain the incidence of voting.

We find that the size of the prize V matters little for the chances of fighting but that there is a strong relationship between λ and the chances of fighting that is well-described by a continuous piecewise linear function with an initially steeply upward sloping segment followed by a relatively flat segment. Near the frontier where $\lambda = 1$ the probability of fighting is small, about 4%. For countries with 20% of frontier income this rises to nearly 14%. For very poor countries with less than 6.5% of frontier GDP this rises somewhat higher to 22%. We also find that fighting does not generally pass a benefit cost test but occurs when there are optimistic beliefs that battle damages being small.

Fighting and λ

Let define $\Xi(x)$ to be the cdf of ϵ. According to the Theorem 1 the probability of fighting is given by $\Xi\left(Q_0\left(\lambda 2r_d/(1-2Q_0)+V\right)\right)$. We first group observations into categories k with cutoff points for λ of the form $(1.25)(1.75)^{k-1}$ and computing for each category the probability of fighting. The maximum value of λ in the data is 110. Table 3 below reports the mean value of λ for each category together with the estimated probability of fighting, the standard error for the binomial average and the number of observations.

λ	Prob. of fighting	SE	Observations
1.08	0.043	0.012	281
1.67	0.085	0.011	634
3.00	0.112	0.012	723
5.05	0.139	0.011	1054
8.58	0.153	0.014	674
15.79	0.196	0.019	420
26.59	0.220	0.025	264
47.74	0.227	0.040	110

Table 3: GDP and the Probability of fighting

As indicated by the standard errors there is not much issue with sampling error. Below in Figure 7 we plot λ against the probability of fighting shown by circles.

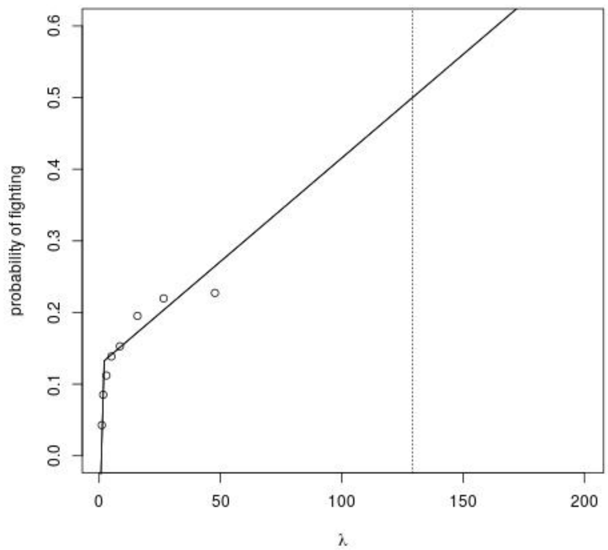

Figure 7: GDP and the Probability of fighting

To a good approximation this is a continuous piecewise linear function with two linear segments. Hence we fitted a segmented linear probability model by choosing the cut point by minimizing the sum of squared residuals as is standard (see, for example, Feder (1975)). The estimating equation is

$$Ey_\tau = \begin{cases} \alpha + \beta_r \lambda_\tau & \lambda_\tau < \lambda_c \\ \alpha + \beta_p \lambda_\tau & \lambda_\tau \geq \lambda_c \end{cases}$$

where τ indexes time and country and y_τ takes on the value one if there is fighting and zero otherwise. The results are below in Table 4 and plotted as the solid line in Figure 7.

			Prior to 1955	
	Coefficient	SE	Coefficient	SE
α	0.133	0.008	0.126	0.012
β_r	0.116	0.036	0.094	0.060
β_p	0.003	0.0005	-0.002	0.003
λ_c	2.09	0.245	2.05	0.417
Observations	4160		1263	

Table 4: Fighting Estimation

The cut point is estimated to lie at 2.09 with corresponding probability 0.13

and the function is estimated to be 0.006 at $\lambda = 1$ and reach 0.5 at $\lambda = 129$ corresponding to the median for $\epsilon = 1$. This is marked with a vertical dotted line.

In Figure 7 the top bin with average $\lambda = 47.74$ has a probability of 0.227 which lies below the estimated line where the fitted probability is 0.265. However there are only 110 observations in the top bin and the standard error is 0.040 so that the discrepancy is less than a standard deviation hence consistent with sampling error.

To check for structural stability we also estimated the model using only data from 1950 and earlier. The results are also shown in Table 4. The coefficient β_p for the upper region, the poor countries, is negative in the earlier sample, but the estimate is very poor and the estimate for the full sample is less than two standard deviations away. The poor estimate in the upper region is due to the fact that there are very few observations of poor countries in the early sample: only 28 observations with $\lambda \geq 12$. The remaining coefficients in the early sample are quite similar to those in the full sample.

Estimating r_d

At $\overline{\lambda} = 129$ there is a 50% chance of fighting, so this is the median. Since the median of ϵ is assume to be one, this means that $Q_0\left(\overline{\lambda}2r_d/(1-2Q_0)+V\right) = 1$. We have already estimated Q_0 to be approximately 0.07. This enables us to establish relationship between r_d, the cost to damage ratio and V the size of the prize. This is

$$r_d = \left(\frac{1-2Q_0}{2\overline{\lambda}}\right)\left(\frac{1}{Q_0} - V\right) = 0.0033\,(14.3 - V)\,.$$

Since $0 \leq V \leq 1$ this give a fairly tight bound on $0.044 \leq r_d \leq 0.048$. This says that paying \$100 results in damage of about \$2,000.

Using the piecewise linear function we see that in the upper region for low income countries

$$\begin{aligned} &\Xi\left(Q_0\left(\lambda 2r_d/(1-2Q_0)+V\right)\right) \\ &= \Xi\left(0.077\lambda + 0.07V\right) \approx 0.124 + 0.0033\lambda, \end{aligned}$$

from which we see that the slope of the cdf in this range is Ξ' = 0.38. Hence the derivative of the probability of civil war with respect to V is 0.026, meaning that an increase in the size of the prize from 0% of GDP to 100% of GDP would increase the chances of fighting in poor countries by only 2.6%.

Implications for V

One implication of this analysis is the the chances of fighting in poor countries are not all that sensitive to V. On the other hand, the intensity of conflict as measured by the cost for each party is $(1-2Q_0)(\rho_f^{-1}/2)V$, which is highly sensitive to V. In other words, the chances of fighting depend on a cost benefit analysis and increasing the size of the prize increases the benefits, but also increases the cost since the cost is endogenous and both parties will incur a greater cost to get a greater prize. Hence there is not so much sensitivity in the cost benefit analysis about whether to engage in civil war. By contrast, if a civil war does start, the intensity of the fighting is directly proportional to the size of the prize. Earlier research did not clearly make this distinction.

Although it matters little in analyzing fighting, the size of the prize does matter for the analysis of voting. One measure of V is discretionary spending in a country with an advanced tax system. In the modern frontier country, the US, for the period 2003-2022 this averaged 7.3% of GDP[5] so we take $V = 0.073$ leading to a corresponding $r_d = 0.047$.

Implications for ϵ

Overall fighting is driven by optimistic draws of ϵ. Recall that the median of ϵ is one and that in this case when the true benefit cost ratio is one there is a 50% of an optimistic shock resulting in a civil war. In fact no countries are so poor that the true benefit cost ratio is greater than or equal to one. In other words, civil wars take place when the true benefit cost ratio is less than one - a civil war has an expected loss - but beliefs about battle damage are optimistic that battle damage is low.

Knowing $r_d = 0.047$ and $V = 0.073$ we can find the implied distribution of ϵ. The bottom of the support of the cdf Ξ is 1.2% meaning that the most optimistic belief is that battle damages are that they are very small. Low values of ϵ below the cutpoint of 2.1% have relatively high density with the probability of being at or below the cutpoint being 13%. Above this the density is lower with the cumulative probability rising to 50% at the median of 100%. As there are no observations above the median there is no information about the distribution of ϵ above this point.

9. Implications for Voting

Having estimated all the relevant parameters we can now make a prediction about how likely countries are to award power through voting. The bottom line is shown in Figure 8 below. The lower curve is the estimated probability of fighting, the same as in Figure 7. The upper curve is the probability that there is no voting: that there is either a dominant party or fighting. The solid part of this curve is derived from the theory and the estimates from the fighting data. The dots represent the actual probabilities using the same bins as in Table 3. The result is striking: the theory fits the data extremely well.

[5]Congressional Budget Office Discretionary Spending in Fiscal Year 2023: An Infographic.

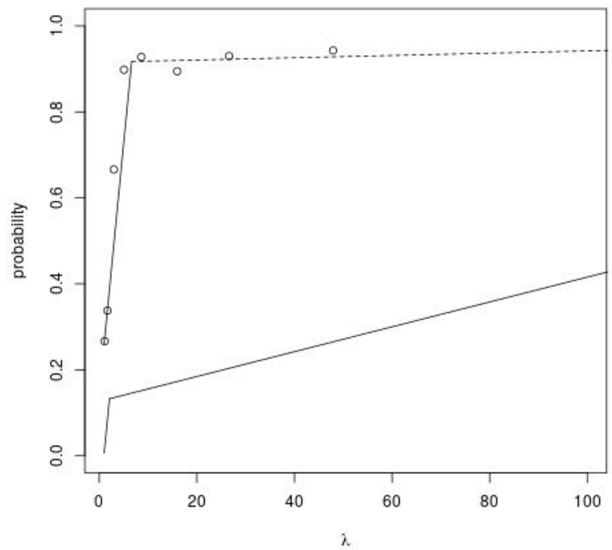

Figure 8: fighting (lower line) and no voting (upper line)

Notice that for poor countries the chances of dominant party rule actually decline with distance to the frontier since the probability of civil war rises faster than the chances of voting decline. For middle income countries the opposite is true with probability of civil war rising slowly but the chances of voting declining rapidly. Note that this is exactly what happens in the data.

Finding Voting

We turn now to the calculations underlying Figure 8. Recall that the probability that there is no voting (either fighting or dominant party) is given from Theorem 1 as

$$\Xi\left[\left(\rho_f^{-1}Q_0+(1-\rho_f^{-1})(1-Q_0)\right)(\lambda 2r_d\rho_f/(1-2Q_0)+V)\right].$$

Using the estimates $Q_0 = 0.07$, $\rho_f = 3.96$, $V = 0.073$ and $r_d = 0.047$ this becomes

$$\Xi\left[0.71\,(0.43\lambda+0.073)\right].$$

From this we can compute that the cut point for no voting is negative so that only the upper segment is relevant. The corresponding probability of no voting in terms of λ is then $0.14+0.12\lambda$. This is the solid upper curve plotted in Figure 8

Voting and Income

As λ increases the probability of no voting approaches 1 and we have no fighting data relevant above the median which for voting is about 3.1. As the data on no voting, indicated by dots using the same bins as for fighting and shown in Figure 8, indicates that the probability can be well approximated by two linear segments we again implemented a segmented regression. We held fixed the curve estimated from the fighting data $0.14 + 0.12\lambda$ up to a cutpoint and using the no voting data above the cutpoint. The results are shown in Table 5 below and is shown by the dashed line in Figure 8.

	Coefficient	SE
α	0.144	N/A
β_r	0.117	N/A
β_p	0.003	0.0009
λ_c	6.63	0.06
Observations	3442	

Table 5: Non-voting estimation

The data is summarized also in Table 6 below.

λ	Predicted	Actual	SE	Observations
1.09	0.272	0.266	0.030	214
1.67	0.339	0.338	0.020	542
2.98	**0.492**	**0.666**	0.020	545
5.06	**0.735**	**0.898**	0.011	916
8.62	0.920	0.928	0.011	582
15.97	0.921	0.895	0.016	361
26.59	0.924	0.930	0.016	257
47.91	0.930	0.942	0.023	105

Table 6: GDP and the Probability of No Voting

Vertical distances are hard to read from the graph in Figure 8. Except for the highlighted cells the values predicted and the actual cell values are quite similar and well within the standard errors of the cell estimates. The exceptions are for the 2.98 and 5.06 cells where the theory underpredicts the data by about 16%. We have enough data that our model - an approximation to the truth - can be rejected by standard statistical tests.

Indeed: the data favors shifting the democracy curve slightly to the left. This poses a problem subsequently when we want to use voting data to assess V on subsamples, that is to estimate equations of the form $Ey_{\ell\tau} = g_\ell(V, \lambda_\tau)$ where $\ell \in \{\text{fight}, \text{novote}\}$. If we estimate this equation on the full sample using the voting data we get the results shown in Table 7 below.

	Vote
V	0.495
SE	0.031
Observations	3422

Table 7: Variable V

As can be seen this leads to a wild overestimate of V, confirming that our approximate model is not "true." Note that if we do the same estimation for the fighting data we naturally recover the value $V = 0.073$. To use the voting data on subsamples we must avoid the overfitting that takes place on the full sample. We do so by adding the difference between the estimated V and 0.073 and estimating $Ey_{\text{novote},\tau} = g_{\text{novote}}(V + (0.495 - 0.073), \lambda_\tau)$ which will result in the proper result of 0.073 when run on the full sample.

10. Income and Democracy

It could be of course that nations are rich because they are democratic, rather than, as in our theory, that they engage in voting because they are rich. There is indeed a long-standing debate in the literature over this. Lipset (1959) argued that democracy is persistent only in rich countries, and this has found support in cross-country regressions such as those of Barro (1999). Acemoglu et al (2008) argue that this correlation is spurious. Although the statistical methods used are controversial (see Che et al (2013)) here we accept the procedure of Acemoglu et al (2008) and show that their evidence that there is "no causal effect of income on democracy" is in fact consistent with our theory that there is.

Cross-sectional Evidence

Acemoglu et al (2008) show that when a measure of democracy is regressed on income and time and country fixed effects are included the correlation between democracy and income is either insignificant or negative depending on the estimation technique. To assess this argument we start by assuming that our model is true and asking what the Acemoglu et al (2008) procedure will find.

We generate an artificial dataset parallel to that used by Acemoglu et al (2008). We take our five year data for their sample period starting in 1955 through 2000 omitting countries for which there are missing per capita GDP observations. We then apply our theory to compute the probability of voting ν_τ : taking the relevant value of λ_τ this is given by

$$E\left[1-\nu_\tau\right] = \begin{cases} 0.144 + 0.116\lambda_\tau & \lambda_\tau < 6.63 \\ 0.912 + 0.000266_\tau\lambda_\tau & \lambda_\tau \geq 6.63 \end{cases}.$$

We presume that ν is what the different democracy indices are trying to measure so take it as our measure of democracy. Letting z denote the logarithm of per capita GDP we then use OLS to estimate the same equation estimated by Acemoglu et al (2008): $\nu_{it} = \alpha\nu_{it-1} + \gamma z_{it-1} + \mu_t + \delta_i$. Below in Table 8 is the result of that estimation along with the findings of Acemoglu et al (2008):

	Theory	Acemoglu et al (2008)	
		OLS	GMM
α	0.429(0.029)	0.379(0.051)	0.489((0.085)
γ	$-0.055(0.004)$	0.010(0.035)	$-0.129(0.076)$

Table 8: Theory versus Data (standard errors in parentheses)

Acemoglu et al (2008) found that when they did OLS using time and country fixed effects the coefficient on z is small and insignificant and when they use more advanced estimation techniques (generalized method of moments) the point estimate in fact becomes negative. If our model is true, and indeed income does cause democracy, in fact the coefficient on z should be negative. Not only that, but the coefficients on both the lag of democracy and the lag on z estimated from our model are quantitatively similar to those coefficients found by Acemoglu et al (2008). The bottom line is that the Acemoglu et al (2008) procedure yields misleading results: it suggests that lagged democracy is important and lagged GDP is not despite the fact that the data is generated without any persistence in democracy and is determined only by income.

Note that the standard errors reported for the theory are the usual OLS standard errors: they are in fact entirely without meaning since the dependent variable is not random. The true standard errors are zero: we can redraw the sample for the same independent variables endlessly and the resulting estimates will always be the same. There is no sampling error, only specification error. This should be a cautionary notes about interpreting computer generated standard errors in the presence of specification error.

To get an intuition into what is going on, observe first that the level of GDP is different than the distance to the frontier but that this does not matter in the presence of time fixed effects since these can be made equal to the log of lagged per capita frontier GDP. However, while the true relationship (even in logs) is highly non-linear rising initially rapidly then becoming extremely flat, the Acemoglu et al (2008) model supposes that the relationship (in logs) is linear.

Suppose, for the purpose of understanding, that nothing changes over time so we are just observing countries some with high and some with low λ. Since the relationship is highly non-linear using z_{it-1} on the right-hand-side leads to a lot of error. On the other hand, country fixed effects can perfectly account for the non-linear relationship, so the coefficient on λ_{it-1} will be zero and the fit will be perfect. In other words: the only way in which the Acemoglu et al (2008) model can accomodate the non-linearity of the data is through the country fixed effects, and it is this specification error that leads to misleading results.

When we account for time this makes things even worse. Poor countries have substantial fluctuations in income over time but because ν is very flat as a function of λ these income fluctuations result in very little change in ν. Hence putting weight on ν leads to substantial error and the regression will avoid doing this.

Does Democracy Cause Income?

The attempt to debunk the idea that income leads to democracy is in part an effort to support the point of view that institutions are crucial and that democracy leads to high income. There is evidence in this direction: see, for example, Madsen, Raschky and Skali (2015), Cervelatti et al (2014) and Acemoglu et al (2019). We do not have a great deal to say about this: it is unlikely that future democracy influences current income which would contradict our model, and we have nothing to say about the frontier itself. Our model is not about why some some nations are richer than others; just why poorer ones are more prone to fighting and dominant party rule.

It is important for our theory, however, that the correlation between distance to the frontier and voting is not driven by more democratic countries catching up to the frontier more quickly. To examine this we looked at detail at countries that have have become substantially richer over time to see if they became rich because of voting. We looked at Korea, Taiwan, Brazil and China. Below are reported in the solid curves the voting index defined as 6 times the ratio of the Polyarchy Index divided by 0.95 times the frontier index with the cutoff at 6. The dotted lines show contemporaneous $\tilde{\lambda}$ and the constant line $\tilde{\lambda} = 4$..

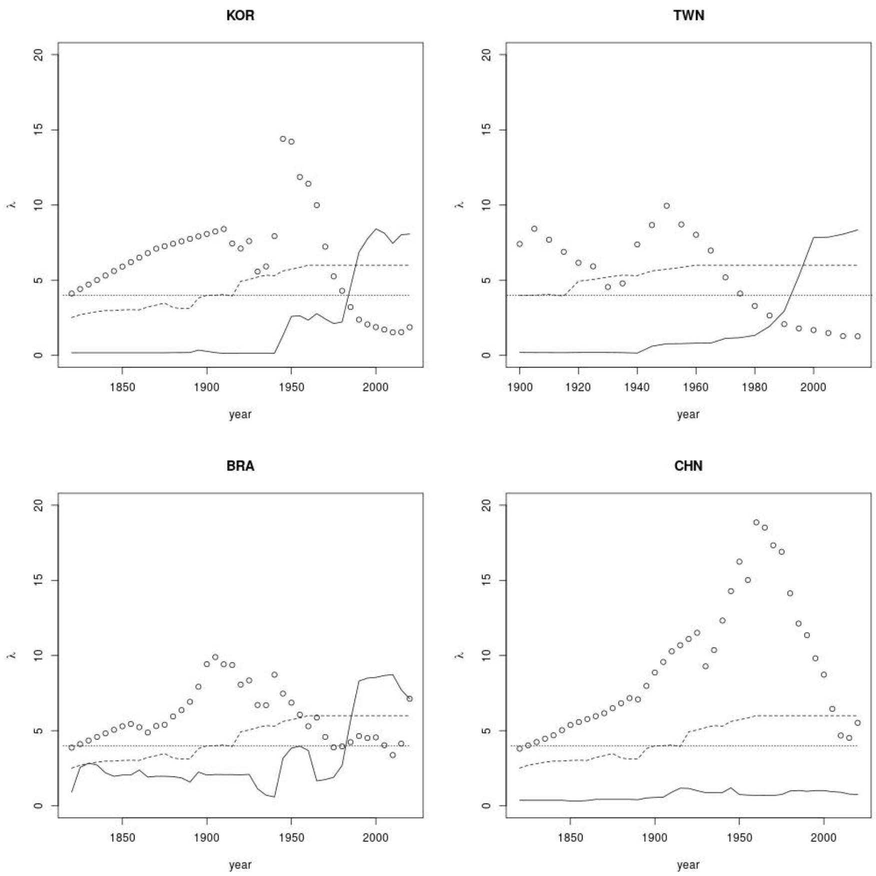

Figure 9: voting (solid line) and λ (dots)

In Korea most of the drop in $\tilde{\lambda}$ occured before the voting index started to increase and $\tilde{\lambda}$ has been quite flat since the advent of voting. The pattern is similar in Taiwan. In Brazil $\tilde{\lambda}$ is actually at at a local minimum when voting starts to rise and is flat after. In China of course there has been no voting but $\tilde{\lambda}$ as we know has fallen enormously under dominant party rule. Note that for both Korea and Brazil the voting index starts to rise around $\tilde{\lambda} = 4$: China is still short of that benchmark.

We examined a number of other countries that did not have growth episodes: Argentina, Australia, Benin, Canada, Chile, France, India, Mexico, South Africa, France, the USA and UK. In no case do we find any evidence that greater voting reduces $\tilde{\lambda}$.

Cross Country Civil War Regressions

Starting with Collier and Hoeffler (1998) the empirical civil war literature has followed a cross-country regression strategy similar to that used in the study of democracy. The point of departure for most modern work is Fearon and Laitin (2003) who regress the incidence of civil war regressed on its lag, on lagged per capita income and other variables. Parallel to the empirical democracy literature, the main finding is that the likelihood of civil war decreases with income. Sylvain and Miguel (2018) point out two main findings: poorer countries are more likely to suffer from civil war, and civil war is more likely to occur when a country is exposed to a negative income shock. The fact that the likelihood of civil war decreass with income overall corresponds to our findings. Unlike the literature on democracy there seems to have been no effort to debunk the major finding by using country fixed effects.

11. Intensity of Conflict and the Prize

Our base assumption is that the size of the prize is constant across countries and times. Since fighting and voting are estimated to be linear except near the single cutpoint it would be enough that the expected size of the prize conditional on λ is constant across countries and times. Here we show that this assumption is in fact consistent with the data.

As indicated earlier changes in the size of the prize V are reflected linearly in the expected per capita efforts $Ef_k = (1 - 2Q_0)(\rho_f^{-1}/2)V/C_o$. While the expected effort is not observed actual per capita effort is. In particular define $\overline{f} = f_i + f_o$ in per capita terms. Then it should be that $E[\overline{f}|\lambda] = E\overline{f}$ independent of λ.. To test if this is the case we can estimate a regression model of the form $E\overline{f}_\tau = \alpha + \beta\lambda_\tau$. Combining the two datasets, for the observations for which data exists we did this with the results show below in Table 9 and with the scatter-plot in Figure 10.

	Coefficient	SE
α	0.0164	.0028
β	0.00026	0.00032
Observations	99	

Table 9: Intensity estimation

As can be seen the slope with respect to λ is slightly positive it is small and if the true slope is zero then there is a 40% probability that such a slope estimate or higher could be generated by sampling error.

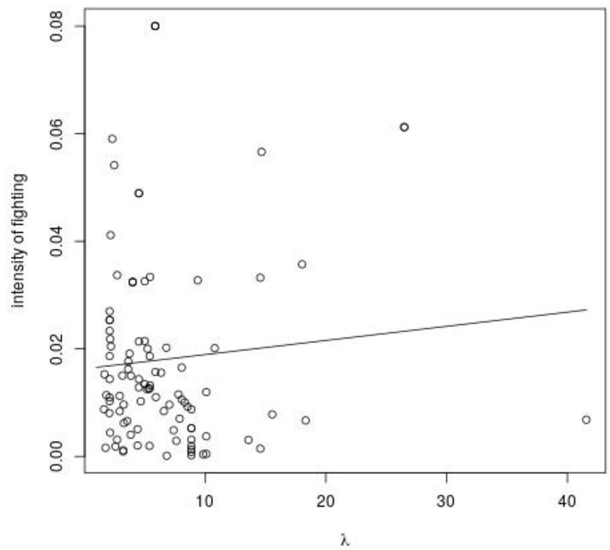

Figure 10: Intensity and Income

The overall conclusion is that there is no convincing evidence that V is correlated with λ. This means that we can legitimately analyze V in subsamples to see if other exogenous variables influence V. Specifically we estimate on subsamples equations of the form $Ey_{\ell\tau} = g_\ell(V, \lambda_\tau)$ where $\ell \in \{\text{fight}, \text{novote}\}$ and g_ℓ is the model with parameters estimated from above. We turn now to this analysis.

Does Oil Matter?

We would expect that oil producing countries would have a larger prize than non-oil producing countries. To assess this we took the subsample of oil producing countries and estimated the size of the prize from the model. As expected it is estimated to be considerably larger for oil producing countries.

We define an oil producing country/period as having oil income at least 5% of GDP. There are 46 countries that met this criterion in at least one period.[6] Conditional on the parameters estimated for the fighting and voting models, we

[6] Angola, Albania, UAR, Argentina, Azerbaijan, Bahrain, Bolivia, Brunei, Canada, Cameroon, Republic of the Congo, Algeria, Ecuador, Egypt, Gabon, Great Britain, Equatorial Guinea, Indonesia, Iran, Iraq, Kazakistan, Kuwait, Libya, Lithuania, Mexico, Malaysia, Nigeria, Netherlands, Norway, Oman, Peru, Qatar, Russia, Saudi Arabia, South Sudan, Senegal, Surinam, Syria, Chad, Turkmenistan, Trinibad and Tobago, Tunisia, Uzbekistan, Venezuela, and Yemen.

use non-linear least squares to estimate a common value of V for these countries. The results are in the first two columns of Table 10 below.

The theory does well: both for fighting and voting the estimated values of V are much larger than the baseline V of 7.3%. Moreover, the estimate of the fight data is well within two standard errors of the estimate from the vote data. While the estimates are both greater than one, the bottom of the two standard deviation confidence interval from the fight data is 0.60 and the bottom for the vote data is 0.66.

Does Polarization Matter?

We would expect that highly polarized countries would have a larger prize than more homogeneous countries. To assess this we used the polarization measure from Esteban, Mayoral and Ray (2012) and took a subsample of countries for which the index is high. We again estimated the size of the prize from the model. We find no evidence it increases either fighting or non-voting.

We define a polarized country/period as having a PEthnoDelta005 index of at least 0.05. There are 44 countries met this criterion in at least one period.[7] Conditional on the parameters estimated for the fighting and voting models we again use non-linear least squares to estimate a common value of V for these countries. The results are in the third and fourth column of Table 10 below.

	Oil		Polarization	
	Fight	Vote	Fight	Vote
V	2.46	1.12	4.60	0.17
SE	0.93	0.23	0.85	0.10
Observations	271	109	393	386

Table 10: Size of the Prize for Oil and for Polarization

Qualitatively the theory does well: both estimates of V are greater than the baseline 0.073. The fighting estimate is much much larger than one, but there is no reason that the consequences of ethnic takeover should be limited by GDP. However, the voting estimate is much smaller than that for the fighting estimate, and indeed the confidence intervals do not overlap. This fact cannot be explained by the simple theory.

The discrepancy between fighting and voting can be from polarization impacting on additional parameters besides the size of the prize. For example, it may be that in highly polarized countries incumbent advantage is smaller because both sides have private armies on a similar footing. Smaller ρ_f does not change fighting but it does reduce the incentive of the incumbent to fight and so

[7] Afghanistan, UAR, Argentina, Azerbaijan, Bulgaria, Bahrain, Bolivia, Republic of the Congo, Cyprus, Dominican Republic, Ecuador, Estonia, Finland, Fiji, Georgia, Guatemala, Hungary, Iran, Iraq, Israel, Kirgizstan, Sri Lanka, Mexico, Myanmar, Mauritania, Mauritius, Malasia, Niger, Nepal, Oman, Panama, Peru, Paraguay, Romania, Russia, South Sudan, Singapore, Surinam, Slovakia, Syria, Trinidad and Tobago, Turkey, and Vietnam.

increases voting. Hence smaller ρ_f would lead to overpredictions of voting which in turn can be explained by larger values of V. Investigating these possibilities is beyond the scope of this paper.

12. Conclusion

We built a theoretical model of the substitutability between voting and fighting incorporating the idea that distance from frontier GDP is crucial in explaining both. We showed that this fits well data on both voting and fighting. In particular, though causality runs from income to democracy adoption in our model, this causality link may be missed by empirical methods that are prominent in the current debate on income and democracy. A key take-away is that reducing inequality through economic development in the sense of increases in productivity should reduce conflict and increase democracy, while reducing inequality through higher prices, subsidies or transfers may have the opposite effect.

References

Acemoglu, D., S. Johnson, J. A. Robinson, J. A. and P. Yared (2008): “Income and Democracy,” *American Economic Review* 98: 808-842.

Acemoglu, D., S. Naidu, P. Restrepo and J. A. Robinson (2019): “Democracy does cause growth,” *Journal of Political Economy* 127: 47-100.

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

Alcalde, J. and M. Dahm (2007): “Tullock and Hirshleifer: a meeting of the minds,” *Review of Economic Design* 11: 101-124.

Ashworth, Scott and Ethan Bueno De Mesquita (2009): “Elections with platform and valence competition,” *Games and Economic Behavior* 67: 191-216.

Barro, Robert J. (1999): “The Determinants of Democracy,” *Journal of Political Economy* 107: S158-S183.

Besley, Timothy J. and Torsten Persson (2008): “The Incidence of Civil War: Theory and Evidence,” NBER.

Bolt, Jutta and Jan Luiten van Zanden (2020): “Maddison style estimates of the evolution of the world economy. A new 2020 update.”

Blattman, C. and E. Miguel (2010): “Civil war,” *Journal of Economic Literature* 48: 3-57.

Cervellati, Matteo, Florian Jung, Uwe Sunde, and Thomas Vischer. "Income and democracy: Comment." American Economic Review 104, no. 2 (2014): 707-719.

Che, Y., Y. Lu, Z. Tao and P. Wang (2013): “The impact of income on democracy revisited,” *Journal of Comparative Economics* 41: 159-169.

Chassang, Sylvain, and Gerard Padró i Miquel. Economic shocks and civil war. SSRN, 2018.

Collier, Paul, and Anke Hoeffler (1998) “On Economic Causes of Civil War,” *Oxford Economic Papers* 50: 563-73.

Dal Bo E. and P. Dal Bo (2011): “Workers, warriors, and criminals: social conflict in general equilibrium,” *Journal of the European Economic Association* 9: 646-677.

Dal Bo, E. and R. Powell (2009): A model of spoils politics. American Journal of Political Science 53: 207-222.

Dixon, Jeffrey S. and Meredith Reid Sarkees (2016): *A GUIDE TO INTRA-STATE WARS: An Examination of Civil, Regional, and Intercommunal Wars, 1816-2014*, Washington, DC: CQ Press 2016.

Dupuy, T. N. (1986): "Mythos or verity? The quantified judgment model and German combat effectiveness," *Military Affairs: The Journal of Military History, Including Theory and Technology*: 204-210.

Ewerhart, C. (2017): "Contests with small noise and the robustness of the all-pay auction," *Games and Economic Behavior* 105: 195-211.

Esteban, J., L. Mayoral and D. Ray (2012): "Ethnicity and conflict: An empirical study," *American Economic Review* 102: 1310-1342.

Fearon, James D. (1995): "Rationalist Explanations for War," *International Organization* 49.

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

Feder, P. I. (1975): "The log likelihood ratio in segmented regression," *The Annals of Statistics* 3: 84-97.

Hillman, A. L. and J. Riley (1989): "Politically Contestable Rents and Transfers," *Economics and Politics* 1: 17-39.

Hirshleifer, J. (1989): "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice* 63: 101-112.

Hirshleifer, J., M. Boldrin and D. K. Levine (2009): "The Slippery Slope of Concession," *Economic Inquiry* 47: 197-205.

Levine, D. K. and A. Mattozzi (2020): "Voter Turnout with Peer Punishment," *American Economic Review* 110: 3298-3314.

Levine, D.K. and A. Mattozzi (2021): "Success in Contests," EUI.

Levine, D. K. (1985): "Sensitivity of MLE to Measurement Error," *Journal of Econometrics* 28: 223-230

Levine, D. K., A. Mattozzi, and S. Modica (2022): *Social Mechanisms and Political Economy: When Lobbyists Succeed, Pollsters Fail and Populists Win*, mimeo EUI.

Levine, D. K. and S. Modica (2018): "Intervention and peace," *Economic Policy* 33: 361–402.

Lipset, Seymour M. (1959): "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review* 53: 69-105.

Madsen, J. B., P. A. Raschky and A. Skali (2015): "Does democracy drive income in the world, 1500–2000?" *European Economic Review* 78: 175-195.

Mawejje, J, McSharry, P. (2021): "The economic cost of conflict: Evidence from South Sudan," *Review of Development Economics* 25: 1969–1990.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer (2024): "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data," V-Dem Workimg Paper.

Ross, Michael and Paasha Mahdavi (2015), "Oil and Gas Data, 1932-2014," Harvard Dataverse.

Tullock, G. (1967): "The welfare costs of tariffs, monopolies, and theft," *Western Economic Journal* 5: 224-232.

Online Appendix 1: Measurement Error

Let $\varphi \geq 1$ be the ratio of combatants and $\tilde{\varphi} = \eta\varphi$ the actual force ratio where the multiplicative shock η is independent of the ratio of combatants φ and η has support bounded above and away from zero. Suppose the support of φ is unbounded above and denote the joint density by $f(\tilde{\varphi}, \varphi)$..

Lemma 1. *The positive correlation conditions*

$$\lim_{\overline{\varphi}\to\infty} \Pr\left(\tilde{\varphi} \leq \sqrt{\overline{\varphi}} | \varphi \geq \overline{\varphi}\right) = 0$$

and

$$\lim_{\overline{\varphi}\to\infty} \frac{\Pr\left(\tilde{\varphi} \leq \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}{\Pr\left(\tilde{\varphi} > \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)} = 0$$

hold.

Proof. For $\eta\varphi \leq \sqrt{\overline{\varphi}}$ to be true it must be true that $\overline{\varphi} \leq \sqrt{\overline{\varphi}}/\underline{\eta}$ which fails as $\overline{\varphi} \to \infty$. Hence the first positive correlation condition is satisfied. Next

$$\lim_{\overline{\varphi}\to\infty} \frac{\Pr\left(\eta\varphi \leq \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}{\Pr\left(\eta\varphi > \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}$$

$$\leq \lim_{\overline{\varphi}\to\infty} \frac{\Pr\left(\underline{\eta}\varphi \leq \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}{\Pr\left(\overline{\eta}\varphi > \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}.$$

For $\overline{\varphi}$ sufficiently large the numerator is 0 and the denominator 1. $\square$

Denote by $0 \leq Q(\tilde{\varphi}) \leq 1/2$ is the probability of success and suppose that $\lim_{\tilde{\varphi}\to\infty} Q(\tilde{\varphi}) = Q_0$. Let $Q_T(\overline{\varphi})$ denote the frequency of success for $\tilde{\varphi} > \overline{\varphi}$ in a sample of size T.

Proposition 1. *There exists a sequence $\overline{\varphi}_T$ such that $Q_T(\overline{\varphi}_T)$ converges in probability to Q_0..*

Proof. For fixed $\overline{\varphi}$ the fact that φ has unbounded support implies that

$$\text{plim}_{T\to\infty} Q_T(\overline{\varphi}) = \frac{\int_{\overline{\varphi}}^{\infty} \int_1^{\infty} Q(\tilde{\varphi}) f(\tilde{\varphi}, \varphi) d\tilde{\varphi} d\varphi}{\int_{\overline{\varphi}}^{\infty} \int_1^{\infty} f(\tilde{\varphi}, \varphi) d\tilde{\varphi} d\varphi}.$$

For any $\overline{\varphi}$ we may write

$$\text{plim}_{T\to\infty} Q_T(\overline{\varphi}) = \frac{\int_{\overline{\varphi}}^{\infty} \int_{\sqrt{\overline{\varphi}}}^{\infty} Q(\tilde{\varphi}) f(\tilde{\varphi}, \varphi) d\tilde{\varphi} d\varphi + \int_{\overline{\varphi}}^{\infty} \int_1^{\sqrt{\overline{\varphi}}} Q(\tilde{\varphi}) f(\tilde{\varphi}, \varphi) d\tilde{\varphi} d\varphi}{\int_{\overline{\varphi}}^{\infty} \int_1^{\infty} f(\tilde{\varphi}, \varphi) d\tilde{\varphi} d\varphi} \equiv \hat{Q}(\overline{\varphi}).$$

We compute the difference between Q_0 and $\hat{Q}(\overline{\varphi})$ in three steps.

First step:

$$\frac{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\sqrt{\overline{\varphi}}} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi} = \Pr\left(\eta\varphi \le \sqrt{\overline{\varphi}}\,|\varphi \ge \overline{\varphi}\right)$$

$$\le \Pr\left(\underline{\eta}\varphi \le \sqrt{\overline{\varphi}}\,|\varphi \ge \overline{\varphi}\right)$$

which is less than $\epsilon/3$ for some large $\overline{\varphi}_{\epsilon}^{1}$ and $\overline{\varphi} > \overline{\varphi}_{\epsilon}^{1}$ by the first correlation condition in Lemma 1.

Second step: For some large $\overline{\varphi}_{\epsilon}^{2}$ and $\overline{\varphi} > \overline{\varphi}_{\epsilon}^{2}$ since $\tilde{\varphi} \ge \underline{\eta}\varphi$ we also have have $|Q(\overline{\varphi}) - Q_0| < \epsilon$.

Third step:

$$\frac{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}$$

$$= \frac{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi + \int_{\overline{\varphi}}^{\infty}\int_{1}^{\sqrt{\varphi_T}} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}$$

$$= \left(1 + \frac{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\sqrt{\overline{\varphi}}} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}\right)^{-1}.$$

We have

$$\frac{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\sqrt{\overline{\varphi}}} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi} = \frac{\Pr\left(\eta\varphi \le \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}{\Pr\left(\eta\varphi > \sqrt{\overline{\varphi}}, \varphi > \overline{\varphi}\right)}$$

and for sufficiently large $\overline{\varphi}_{\epsilon}^{3}$ and $\overline{\varphi} > \overline{\varphi}_{\epsilon}^{3}$ this goes to zero by the second correlation condition in Lemma 1. Hence

$$\frac{\int_{\overline{\varphi}}^{\infty}\int_{\sqrt{\overline{\varphi}}}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi}{\int_{\overline{\varphi}}^{\infty}\int_{1}^{\infty} f(\tilde{\varphi},\varphi)d\tilde{\varphi}d\varphi} \to 1.$$

Taking $\overline{\varphi}_{\epsilon} = \overline{\varphi}_{\epsilon}^{1} + \overline{\varphi}_{\epsilon}^{2} + \overline{\varphi}_{\epsilon}^{3}$ we see that for $\overline{\varphi} \ge \overline{\varphi}_{\epsilon}$we have $|\hat{Q}(\overline{\varphi}) - Q_0| \le \epsilon$. For each ϵ choose T such that $\Pr\left(|Q_T(\overline{\varphi}_{\epsilon}) - \hat{Q}(\overline{\varphi})| > \epsilon\right) < \epsilon$ and define $\overline{\varphi}_T = \overline{\varphi}_{\epsilon}$.. Letting $\epsilon \to 0$ now gives the desired result. $\square$