14.773 Political Economy of Institutions and Development.
Lectures 6 and 7: Weak States and State Building

Daron Acemoglu

MIT

February 25 and 27, 2014.

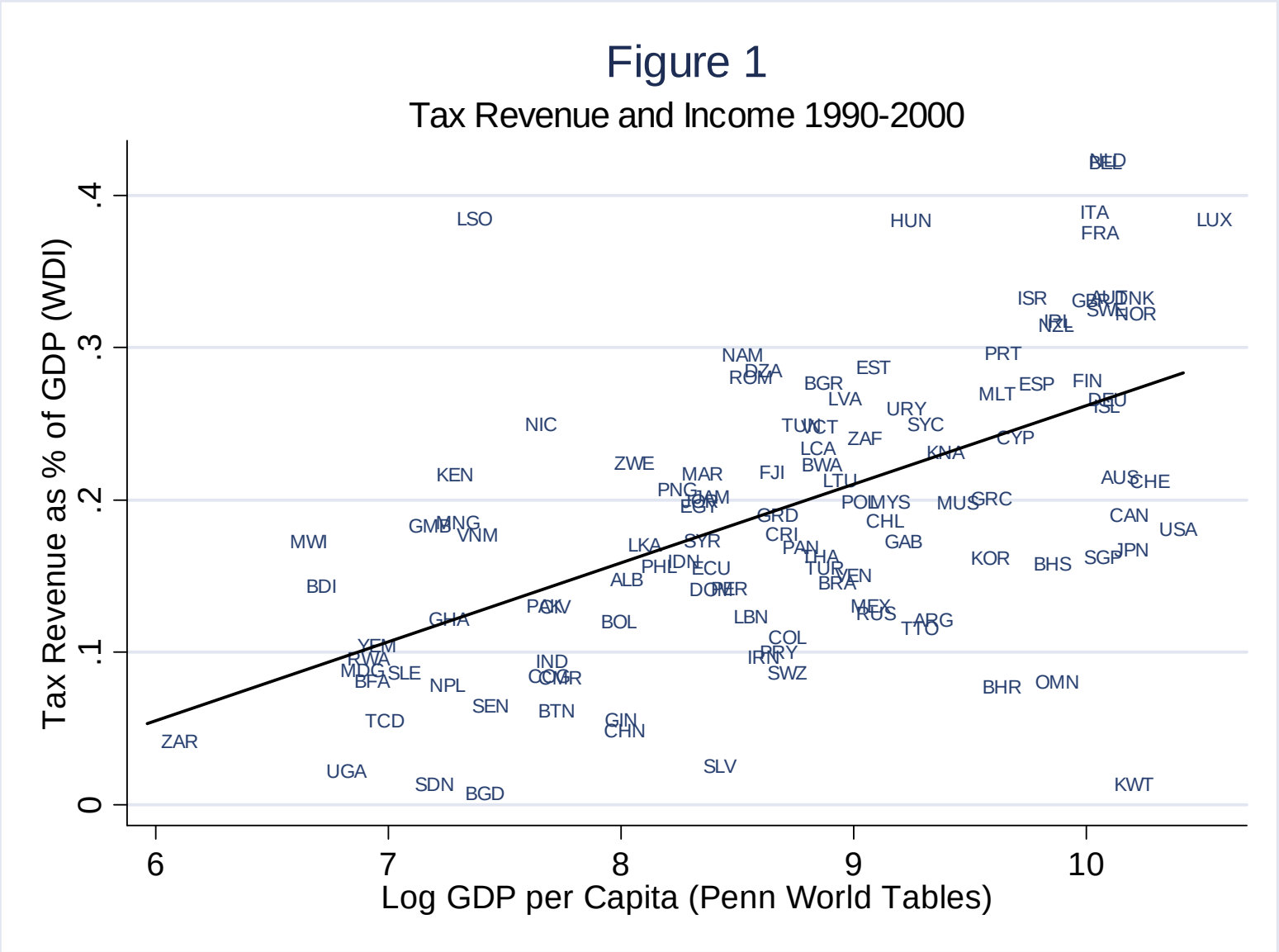
Introduction

- A simple reading of the model/ideas so far would suggest that when states have less power to tax or interfere in economic activities, then there will be fewer political economy distortions and better economic outcomes.
- However, as we will next see, "weak states" are generally associated with worse economic outcomes.
- In fact, many of the ideas we have seen so far might have little relevance to the problem of economic development in some parts of the world where the state is notable in its absence.
- In Joel Migdal's words in *Strong Societies and Weak States*:

  "In parts of the Third World, the inability of state leaders to achieve predominance in large areas of their countries has been striking..."

- In traditional political science, much emphasis on "state capacity" and "weak states".

Income and Taxes



Other Evidence

- Bocktstette, Chanda, and Putterman (2002): countries with early state formation grow faster (but this work should be read with some caution, since they are not richer today according to their empirical work....).
- Gennaioli and Rainer (2007): within Africa, countries with a history of centralized tribal institutions is associated with higher program of public goods.
- Michalopoulos and Papaioannou (2013): within Africa, ethnic groups with the history of centralized tribal institutions are richer.

State Centralization within Uganda

Map 1 Distribution of centralized and fragmented ethnic groups across Uganda regions

Public Goods and State Centralization within Uganda

Table 1 Precolonial centralization and public goods in Uganda

Region Precolonial institutions of ethnic groups	Central Centr	Western Centr	Eastern Mixed	Northern Fragm
% of roads paved in 2002	13.37	10.32	10.89	1.33
Infant mortality in 2001	71.9	97.8	89.3	105.9
% of children under five years with diarrhoea in 2001	14.5	16	23.3	26.7
Availability of sewerage system in 2000 (% of households)	15	14	9	6
Piped water inside house in 2000 (% of households)	10	10	8	5
Availability of latrine or human waste disposal service in 2000 (% of households)	96	86	77	67
Adult literacy rate in 1997	72	61	54	54
Adequacy of facility & equipment at primary schools in 2000 (% of households satisfied)	62	72	55	51

Public Goods and State Centralization across Sub-Saharan Africa

Table 2 Precolonial centralization and public goods provision

	% of roads paved in 1990–2000		% of infants immunized for DPT in 2001		Infant mortality in 1960–2001		Adult illiteracy rate in 1970–2002		School attainment in 1960–1990	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Centralization	21.02***	21.53***	36.79***	36.08***	−35.24**	−42.5***	−18.74**	−23.77***	1.24**	1.24**
	(7.21)	(7.04)	(6.53)	(7.13)	(14.79)	(15.12)	(9.04)	(8.29)	(0.47)	(0.49)
Log of initial GDP/cap		4.95		0.9		−23.8**		−11.17**		0.36
		(3.38)		(2.5)		(9.26)		(4.56)		(0.61)
Constant	7.12**	−26.72	38.11***	33.22*	146.6***	306.72***	66.94***	145.63***	1.2***	−1.12
	(2.65)	(22.73)	(4.85)	(17.37)	(10.01)	(63.48)	(5.95)	(32.36)	(0.33)	(3.91)
Obs.	40	40	42	41	42	40	37	36	26	26
Rsq.	0.24	0.27	0.33	0.31	0.13	0.23	0.1	0.26	0.12	0.14

Prosperity and State Centralization within Countries

PRE-COLONIAL ETHNIC INSTITUTIONS AND REGIONAL DEVELOPMENT WITHIN AFRICAN COUNTRIES[a]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Pre-Colonial Ethnic Institutions and Regional Development Within African Countries												
	All Observations											
Jurisdictional Hierarchy	0.3260***	0.2794***	0.2105***	0.1766***								
	(0.0851)	(0.0852)	(0.0553)	(0.0501)								
Binary Political Centralization					0.5264***	0.5049***	0.3413***	0.3086***				
					(0.1489)	(0.1573)	(0.0896)	(0.0972)				
Petty Chiefdoms									0.1538	0.1442	0.1815	0.1361
									(0.2105)	(0.1736)	(0.1540)	(0.1216)
Paramount Chiefdoms									0.4258*	0.4914*	0.3700**	0.3384**
									(0.2428)	(0.2537)	(0.1625)	(0.1610)
Pre-Colonial States									1.1443***	0.8637***	0.6809***	0.5410***
									(0.2757)	(0.2441)	(0.1638)	(0.1484)
Adjusted R-squared	0.409	0.540	0.400	0.537	0.597	0.661	0.593	0.659	0.413	0.541	0.597	0.661
Observations	682	682	682	682	682	682	682	682	682	682	682	682
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

State Capacity

- By weak states, we mean states that lack "state capacity".
- But then, what is state capacity?
- Four different aspects (mostly interwoven).
  1. Max Weber's monopoly of legitimate violence so as to enforce law and order and eliminate competitors.
  2. Ability to tax and regulate economic activity (related, but of course not identical, to the share of tax revenue shown above).
  3. Infrastructural power/capacity of the state—related to the presence of the state and its functionaries.
  4. Max Weber's rational/autonomous bureaucracy—related the ability of state institutions to be somewhat autonomous from politically powerful groups in society.

Key Political Economy Question

- **Key political economy question** that is rarely asked: why are states weak?
- In other words, why do states remain weak, especially if:
  - state weakness is economically costly;
  - most political powerful groups and individuals would prefer to control a strong state.

Plan

- We will start with a simple model of weak vs. strong states based on ability to tax and regulate economic activity.
- We will then turn to models and empirical evidence on different aspects of the state.
- Throughout the emphasis will be on:
  - why state weakness affects economic (and sometimes political) outcomes;
  - why state weakness emerges as equilibrium;
  - how state building takes place.

Environment

- Time is discrete and indexed by t.
- There is a set of citizens, with mass normalized to 1, and a ruler.
- All agents discount the future with the discount factor β, and have the utility function

$$u_t = \sum_{j=0}^{\infty} \beta^j \left[c_{t+j} - e_{t+j} \right],$$

where c_{t+j} is consumption and e_{t+j} is investment (effort), and we assume that the ruler incurs no effort cost.

- Each citizen i has access to the following Cobb-Douglas production technology to produce the unique final good in this economy:

$$y_t^i = \frac{1}{1-\alpha} A_t^{\alpha} \left(e_t^i \right)^{1-\alpha},$$

where A_t denotes the level of public goods (e.g., the state of the infrastructure, or the degree of law and contract enforcement between private citizens), at time t.

Environment (continued)

- The level of A_t will be determined by the investment of the ruler
  - a certain degree of state investment in public goods, the infrastructure or law-enforcement is necessary for production;
  - in fact, investment by the state is complementary to the investments of the citizens.
- The ruler sets a tax rate τ_t on income at time t.
- Each citizen can decide to hide a fraction z_t^i of his output, which is not taxable, but hiding output is costly, so a fraction δ of it is lost in the process.
- This formulation with an economic exit option for the citizens is a convenient, though reduced-form, starting point.
- Given a tax rate τ_t, the consumption of agent i is:

$$c_t^i \leq \left[(1-\tau_t)\left(1-z_t^i\right) + (1-\delta)\, z_t^i \right] y_t^i,$$

where tax revenues are

$$T_t = \tau_t \int \left(1-z_t^i\right) y_t^i di. \tag{1}$$

Environment (continued)

- The ruler at time t decides how much to spend on A_{t+1}, with production function

$$A_{t+1} = \left[\frac{(1-\alpha)\phi}{\alpha} G_t\right]^{1/\phi} \tag{2}$$

where G_t denotes government spending on public goods, and $\phi > 1$, so that there are decreasing returns in the investment technology of the ruler (a greater ϕ corresponds to greater decreasing returns).

- The term $\left[(1-\alpha)\phi/\alpha\right]^{1/\phi}$ is included as a convenient normalization.
- In addition, (2) implies full depreciation of A_t, which simplifies the analysis below.
- The consumption of the ruler is whatever is left over from tax revenues after his expenditure and transfers,

$$c_t^R = T_t - G_t.$$

Timing of Events

- The economy inherits A_t from government spending at time $t-1$.
- Citizens choose their investments, $\left\{e_t^i\right\}$.
- The ruler decides how much to spend on next period's public goods, G_t, and sets the tax rate τ_t.
- Citizens decide how much of their output to hide, $\left\{z_t^i\right\}$.

First-Best Allocation

- The first best allocation maximizes net output:
- Given by public goods investment

$$A_t = \beta^{1/(\phi-1)}$$

and

$$e_t^{fb} = \beta^{1/(\phi-1)} \text{ and } y_t^{fb} = \frac{1}{1-\alpha}\beta^{1/(\phi-1)}.$$

Markov Perfect Equilibrium

- Exit options:

$$z_t^i \begin{cases} = 1 & \text{if } \tau_t > \delta \\ \in [0,1] & \text{if } \tau_t = \delta \\ = 0 & \text{if } \tau_t < \delta \end{cases}.$$

Then, the optimal tax rate for the ruler is

$$\tau_t = \delta. \tag{3}$$

- Next, investment decisions:

$$e_t^i = (1-\delta)^{1/\alpha} A_t. \tag{4}$$

- Substituting (3) and (4) into (1), the equilibrium tax revenue as a function of the level of infrastructure is

$$T(A_t) = \delta y_t = \frac{(1-\delta)^{(1-\alpha)/\alpha} \delta A_t}{1-\alpha}. \tag{5}$$

Markov Perfect Equilibrium (continued)

- The ruler will choose public investment, G_t to maximize his net present value, written recursively as:

$$V(A_t) = \max_{A_{t+1}} \left\{ T(A_t) - \frac{\alpha}{(1-\alpha)\phi} A_{t+1}^{\phi} + \beta V(A_{t+1}) \right\}$$

- First-order condition for the ruler:

$$\frac{\alpha}{1-\alpha} A_{t+1}^{\phi-1} = \beta V'(A_{t+1}).$$

- The marginal cost of greater investment in infrastructure for next period must be equal to to the greater value that will follow from this.
- The envelope condition:

$$V'(A_t) = T'(A_t) = \frac{(1-\delta)^{(1-\alpha)/\alpha}\delta}{1-\alpha}. \tag{6}$$

- The value of better infrastructure for the ruler is the additional tax revenue that this will generate, which is given by the expression in (6).

Markov Perfect Equilibrium (continued)

- Equilibrium actions of the ruler are:

$$A_{t+1} = A[\delta] \equiv \left(\frac{\beta (1-\delta)^{\frac{1-\alpha}{\alpha}} \delta}{\alpha} \right)^{\frac{1}{\phi-1}} \text{ and } G_t = \frac{\alpha}{(1-\alpha)\phi} (A[\delta])^{\phi}, \tag{7}$$

- And therefore:

$$V^*(A_t) = \frac{(1-\delta)^{(1-\alpha)/\alpha} \delta A_t}{1-\alpha} + \frac{\beta(\phi-1)(1-\delta)^{(1-\alpha)/\alpha} \delta}{(1-\beta)(1-\alpha)\phi} A[\delta].$$

Markov Perfect Equilibrium (continued)

- Summarizing:

Proposition: There exists a unique MPE where, for all t, $\tau_t(A_t) = \delta$, $G(A_t)$ is given by (7), and, for all i and t, $z^i(A_t) = 0$ and $e^i(A_t)$ is given by (4). The equilibrium level of aggregate output is:

$$Y_t = \frac{1}{1-\alpha}(1-\delta)^{(1-\alpha)/\alpha} A[\delta]$$

for all $t > 0$ and

$$Y_0(A_0) = \frac{1}{1-\alpha}(1-\delta)^{(1-\alpha)/\alpha} A_0.$$

Second Best

- What is the level of δ—economic strength of the state—that maximizes output.
- Considered a problem

$$\max_{\delta} Y_t(\delta) = \frac{1}{1-\alpha}(1-\delta)^{(1-\alpha)/\alpha} A[\delta],$$

where $A[\delta]$ is given by (7).

- The output maximizing level of the economic power of the state, denoted δ^*, is

$$\delta^* = \frac{\alpha}{\phi(1-\alpha)+\alpha}. \tag{8}$$

Second Best (continued)

- If the economic power of the state is greater than δ^*, then the state is too powerful, and taxes are too high relative to the output-maximizing benchmark.
  - This corresponds to the standard case that the political economy literature has focused on.
- In contrast, if the economic power of the state is less than δ^*, then the state is not powerful enough for there to be sufficient rents in the future to entice the ruler to invest in public goods (or in the infrastructure, law-enforcement etc.).
- This corresponds to the case of "weak states".
  - With only limited power of the state to raise taxes in the future, the ruler has no interest in increasing the future productive capacity of the economy.

Political Power of the State

- Do the same insights applied to the political power of the state?
- Generally yes,

Extended Environment

- Citizens decide replacement: $R_t \in \{0, 1\}$.
- After replacement, the existing ruler receives 0 utility, and citizens reclaim a fraction η of the tax revenue and redistribute it to themselves as a lump sum transfer, S_t.
- Replacement is costly: the cost of replacing the current ruler with a new ruler equal to $\theta_t A_t$, where θ_t is a nonnegative random variable with a continuous distribution function $\tilde{F}_\lambda$, with (finite) density $\tilde{f}_\lambda$.
- Assume that

$$\frac{\tilde{f}_\lambda(x)}{1-\tilde{F}_\lambda(x)} \text{ is nondecreasing in } x \text{ and } \tilde{F}_\lambda(0) < 1, \qquad \text{(A1)}$$

which is the standard monotone hazard (or log concavity) assumption.

Timing of Events

- The economy inherits A_t from government spending at time $t-1$.
- Citizens choose their investments, $\left\{e_t^i\right\}$.
- The ruler decides how much to spend on next period's public goods, G_t, and sets the tax rate τ_t.
- Citizens decide how much of their output to hide, $\left\{z_t^i\right\}$.
- θ_t is realized.

Citizens choose R_t. If $R_t = 1$, the current ruler is replaced and the tax revenue is redistributed to the citizens as a lump-sum subsidy $S_t = \eta T_t$.

Markov Perfect Equilibrium

- Suppose

$$\delta \in (\delta^*, \alpha), \tag{A2}$$

  where δ^* is given by (8).
- This assumption ensures that taxes are always less than the value α that maximizes ruler utility, and also allows the potential for excessively high taxes (i.e., $\tau > \delta^*$).
- Citizens will replace the ruler, i.e., $R_t = 1$, whenever

$$\theta_t < \frac{\eta T_t}{A_t}. \tag{9}$$

- Therefore, the probability that the ruler will be replaced is $\tilde{F}_\lambda(\eta T_t / A_t)$.

Markov Perfect Equilibrium (continued)

- To simplify the notation, define

$$\mathcal{T}(\tau_t) = \frac{(1-\tau_t)^{(1-\alpha)/\alpha}\tau_t}{1-\alpha}.$$

- Also parameterize $\tilde{F}_\lambda(x/\eta) = \lambda F(x)$ for some continuous distribution function F with (finite) density f. Then

$$\begin{aligned} V(A_t) \quad = \quad & \\ \max_{\tau_t \in [0,\delta], A_{t+1}} \quad & \left\{ \left(1-\lambda F\left(\mathcal{T}(\tau_t)\right)\right) \left(\mathcal{T}(\tau_t) A_t - \frac{\alpha}{\phi(1-\alpha)} A_{t+1}^{\phi} \right) \right. \\ & \left. + \beta \left(1-\lambda F\left(\mathcal{T}(\tau_t)\right)\right) V(A_{t+1}) \right\}. \end{aligned}$$

- Now the ruler's maximization problem involves two choices, τ_t and A_{t+1}, since taxes are no longer automatically equal to the maximum, δ.

Markov Perfect Equilibrium (continued)

- In this choice, the ruler takes into account that a higher tax rate will increase the probability of replacement.
- The first-order condition with respect to τ_t yields:

$$\frac{\partial \mathcal{T}(\tau_t)}{\partial \tau_t} \times [(1 - \lambda F(\mathcal{T}(\tau_t))) - \lambda f(\mathcal{T}(\tau_t)) \left(\mathcal{T}(\tau_t) - \frac{G_t}{A_t} + \beta \frac{V(A_{t+1})}{A_t} \right)] \geq 0,$$

  and $\tau_t \leq \delta$ with complementary slackness
- The envelope condition is now

$$V'(A_{t+1}) = (1 - \lambda F(\mathcal{T}(\tau_{t+1}))) \mathcal{T}(\tau_{t+1}). \quad (10)$$

- It only differs from the corresponding condition above, (6), because with probability $\lambda F(\mathcal{T}(\tau_{t+1}))$, the ruler will be replaced and will not enjoy the increase in future tax revenues.

Markov Perfect Equilibrium (continued)

- Using this, the first-order condition with respect to A_{t+1} implies that in an interior equilibrium:

$$A_{t+1} = A\left[\tau_{t+1}\right] \equiv \left(\alpha^{-1}\beta\left(1-\lambda F\left(\mathcal{T}\left(\tau_{t+1}\right)\right)\right)\left(1-\tau_{t+1}\right)^{\frac{1-\alpha}{\alpha}}\tau_{t+1}\right)^{\frac{1}{\phi-1}}$$

- The optimal value of A_{t+1} for the ruler depends on τ_{t+1} since, from the envelope condition, (10), the benefits from a higher level of public good are related to future taxes.
- Also suppose:

$$\left(1-\frac{\beta}{\phi}\left(1-\lambda F\left(0\right)\right)\right)^2-\left(\phi-1\right)\frac{\beta}{\phi}\left(1-\lambda F\left(0\right)\right)>0. \qquad \text{(A3)}$$

- This assumption requires $\beta\left(1-\lambda F\left(0\right)\right)$ not to be too large, and can be satisfied either if β is not too close to 1 or if $\lambda F\left(0\right)$ is not equal to zero.

Markov Perfect Equilibrium (continued)

- Then we have:

Proposition: Suppose (A1), (A2) and (A3) hold. Then, in the endogenous replacement game of this section, there exists a unique steady-state MPE. In this equilibrium, there exists $\lambda^* \in (0, \infty)$ such that output is maximized when $\lambda = \lambda^*$.

Markov Perfect Equilibrium (continued)

- Similar to the case of the economic power of the state, there is an optimal level of the political power of the state.
- Intuitively, when $\lambda < \lambda^*$, the state is too powerful and taxes are too high and citizens' investments are too low.
- When $\lambda > \lambda^*$, the state is too weak and taxes and public investments are too low.
- The intuition is also related to the earlier result.
- When the state is excessively powerful, i.e., $\lambda < \lambda^*$, citizens expect high taxes and choose very low levels of investment (effort).
- In contrast, when $\lambda > \lambda^*$, the state is excessively weak and there is the reverse holdup problem; high taxes will encourage citizens to replace the ruler, and anticipating this, the ruler has little incentive to invest in public goods, because he will not be able to recoup the costs of current investment in public goods with future revenues.

Consensually Strong States

- Neither the analysis of the economic or the political power of the state generate a pattern in which better institutional controls lead to greater government spending.
- But comparison of OECD to Africa might suggest such a pattern.
- Why would this be the case?
- One possibility: go beyond MPE
- *Consensually Strong States*: citizens have low costs of replacing governments, a new look at SPE, where if the government does not follow citizens' wishes, it is replaced.
- Consensually Strong States can generate the pattern of greater public good provision in situations of better controls on government.

Ideas about State Building

- Otto Hinze and later Charles Tilly emphasized the role of inter-state wars in the formation of the state.
- Tilly:

> *"War made the state, and state made war."*

- Based on this, Jeffrey Herbst in *States and Power in Africa* suggested that the weakness of the sub-Saharan African states is due to its difficult terrain and low population density that discouraged inter-state warfare.

A Simple Model of State Building

- Besley and Persson (2009) provided a simple formalization of Tilly.
- "State power" is a state variable (i.e., is persistent "stock"). It enables more efficient taxation.
- In an economy with two competing groups, the group in power may not want to invest in state power if it expects to lose power because then this power will be in the hands of its rival group.
- However, if there is a threat of war, then building state power becomes a necessity.

Evidence

Table 2: Economic and Political Determinants of Fiscal Capacity

	(1) One Minus Share of Trade Taxes in Total Taxes	(2) One Minus Share of Trade and Indirect Taxes in Total Taxes	(3) Share of Income Taxes in GDP	(4) Share of Taxes in GDP
Incidence of External Conflict up to 1975	0.921***	0.683***	0.747***	0.678***
	(0.229)	(0.201)	(0.246)	(0.211)
Incidence of Democracy up to 1975	0.005	– 0.037	0.057	0.097
	(0.085)	(0.096)	(0.062)	(0.064)
Incidence of Parliamentary Democracy up to 1975	0.123	0.208**	0.231***	0.166**
	(0.086)	(0.094)	(0.074)	(0.069)
English Legal Origin	– 0.013	- 0.012	– 0.015	0.013
	(0.069)	(0.061)	(0.056)	(0.051)
Socialist Legal Origin	0.051	– 0.332***	– 0.155**	– 0.110
	(0.095)	(0.084)	(0.065)	(0.082)
German Legal Origin	0.283***	0.290***	0.295***	0.206***
	(0.064)	(0.093)	(0.084)	(0.065)
Scandinavian Legal Origin	0.333***	0.195**	0.364**	0.363***
	(0.068)	(0.078)	(0.141)	(0.092)
Observations	104	104	104	104
R-squared	0.412	0.435	0.628	0.639

Robust standard errors in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%
All specifications include regional fixed effects (for eight regions).

But

- Osafo-Kwaako and Robinson (2013), using the Standard Cross-Cultural Sample, show that these ideas have limited explanatory power for Africa.

Table 3
Correlates of state formation (full SCCS sample).

	Dependent variable: jurisdictional hierarchy beyond local community (SCCS variable §237)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population density	0.386***	0.363***					0.361***	0.295***
	(0.0407)	(0.0403)					(0.0609)	(0.0590)
Africa * Population density	−0.317***	−0.325***					−0.217	−0.173
	(0.115)	(0.112)					(0.180)	(0.171)
Being attacked			0.293*	0.364**			0.263	0.253
			(0.168)	(0.146)			(0.170)	(0.175)
Africa * Being attacked			−0.315	−0.472			−0.102	−0.220
			(0.382)	(0.379)			(0.366)	(0.382)
Trade					0.191*	0.172	0.0113	−0.0375
					(0.112)	(0.115)	(0.107)	(0.117)
Africa * Trade					−0.000762	0.104	0.294	0.408*
					(0.258)	(0.263)	(0.227)	(0.227)
Africa dummy	1.631***	1.644***	0.975	1.064	0.863	0.253	0.595	0.263
	(0.442)	(0.413)	(0.780)	(0.793)	(0.823)	(0.862)	(1.271)	(1.326)
F-stat	0.41	0.13	0.00	0.10	0.67	1.37	1.06	1.46
p-Value	0.52	0.72	0.95	0.76	0.41	0.24	0.37	0.23
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes
Disease environment control	No	Yes	No	Yes	No	Yes	No	Yes
Observations	182	182	152	152	98	98	94	94
R-squared	0.342	0.369	0.040	0.206	0.099	0.272	0.381	0.440

But (continued)

Table 4
Correlates of state formation.

	Dependent variable: levels of sovereignty (SCCS variable §83)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population density	0.388***	0.379***					0.397***	0.325***
	(0.0353)	(0.0379)					(0.0542)	(0.0516)
Africa * Population density	−0.288**	−0.314**					−0.293	−0.272
	(0.131)	(0.129)					(0.220)	(0.199)
Being attacked			0.199	0.263*			0.106	0.0423
			(0.156)	(0.140)			(0.147)	(0.137)
Africa * Being attacked			−0.645*	−0.789**			−0.263	−0.397
			(0.347)	(0.327)			(0.323)	(0.331)
Trade					0.248**	0.260**	0.0255	0.0102
					(0.103)	(0.104)	(0.0898)	(0.0918)
Africa * Trade					−0.254	−0.165	0.0511	0.160
					(0.340)	(0.362)	(0.375)	(0.393)
Africa dummy	1.415**	1.445***	1.428**	1.456**	1.695	0.980	1.995	1.620
	(0.560)	(0.535)	(0.709)	(0.694)	(1.126)	(1.231)	(1.569)	(1.664)
F-stat	0.62	0.27	2.07	3.10	0.00	0.08	0.24	0.58
p-Value	0.43	0.60	0.15	0.08	0.99	0.78	0.87	0.63
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes
Disease environment control	No	Yes	No	Yes	No	Yes	No	Yes
Observations	182	182	152	152	99	99	95	95
R-squared	0.355	0.373	0.028	0.187	0.108	0.321	0.401	0.505

Stationary Bandits

- Mancur Olson, also based on some of Tilly's ideas, argued that the origins of the state lie in organized banditry (e.g., Olson, 1993, and McGuire and Olson, 1996).
- A roving bandit will apply maximal extraction (as in the model above in the MPE with δ high). Roving bandits arise when the bandits themselves don't have any security and for this or other reasons have a short horizon.
- A stationary bandit, with a longer horizon, will act like a state, encouraging production and taking more moderate taxes (as in the SPE of the models we have seen before).
- A stationary bandit ultimately becomes a state.

Evidence

- Sanchez de la Sierra (2014) provides evidence for this perspective by exploiting the differential increases in incentives of armed groups in the civil war of Eastern Congo to become "stationary bandits" because of the Colton price hike.
- Colton is easier to tax because it's much harder to conceal than gold, so he uses gold as a control.
- He therefore hypothesizes that "attempted conquests" should increase due to the higher interaction of Colton deposits and Colton price, but not the same for gold.
- But does this have anything to do with State building?

Evidence (continued)

Table 3: Effects of price shocks, conquests

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coltan X Coltan Price	0.25*	0.61***	0.26*	0.30**	0.60**	0.59**	0.60**	0.30**	0.66***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.24)	(0.24)	(0.25)	(0.14)	(0.20)
Coltan X Coltan Price X Far		-0.58***			-0.57*	-0.56*	-0.57*		-0.66**
		(0.19)			(0.29)	(0.29)	(0.29)		(0.27)
Gold X Gold Price	-0.01***	0.00	-0.01***	-0.02	-0.02	-0.02	-0.03		
	(0.00)	(0.01)	(0.00)	(0.02)	(0.03)	(0.03)	(0.04)		
Gold X Gold Price X Far		-0.01			0.01	0.02	0.02		
		(0.01)			(0.05)	(0.04)	(0.04)		
Congolese Army, Village			-0.06**						
			(0.03)						
Constant	0.12***	-70.08	0.12***	0.15*	-0.02	-0.04	-0.05	0.09***	0.09***
	(0.03)	(74.18)	(0.03)	(0.09)	(0.09)	(0.08)	(0.09)	(0.02)	(0.02)
Observations	1,342	1,342	1,342	732	732	678	732	244	244
R-squared	0.21	0.33	0.22	0.34	0.35	0.29	0.37	0.68	0.69
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Village FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Region*Year FE	NO	NO	NO	NO	NO	NO	YES	NO	NO
Distance to road	NO	NO	NO	NO	NO	YES	NO	NO	NO
Sample	1998-2009	1998-2009	1998-2009	1998-2003	1998-2003	1998-2003	1998-2003	1999-2000	1999-2000

But

- Empirically, long-lived dictators (Mobutu, Mugabe, the Duvaliers in Haiti) are not more developmental, and if anything seem to be among the most kleptocratic.
- Conceptually, this equates the state with organized banditry. But is that right?
- Theoretically, Olson's vision is too narrow also.
  - Acemoglu and Robinson (APSR 2006): the relationship between entrenchment and likelihood to take actions against economic development is inverse U-shaped. This is because a very non-entrenched dictator has no reason to sabotage development in order to save his future rents.
  - Acemoglu, Golosov and Tsyvinski (JET 2010): from a repeated games perspective, a less entrenched dictator may be easier to discipline. This is because if he deviates, society can more easily punish him by removing him from power. We next explain this result.

Power and Efficiency

- Consider an infinite horizon economy in discrete time with a unique final good, consisting of N parties (groups).
- Each party j has utility at time $t = 0$ given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u_j(c_{j,t}, l_{j,t}),$$

where $c_{j,t}$ is consumption, $l_{j,t}$ is labor supply (or other types of productive effort), and $\mathbb{E}_0$ denotes the expectations operator at time $t = 0$.

- Labor supply lies in $[0, \bar{l}]$ for each party, and let us make the usual assumptions on the utility function.
- Aggregate output is given by

$$Y_t = \sum_{j=1}^{N} l_{j,t}.$$

First Best

- First best is straightforward to define.
- Let us introduce the *Pareto weights* vector denoted by $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$, where $\alpha_j \geq 0$ for $j = 1, ..., N$ denotes the weight given to party j, with $\sum_{j=1}^{N} \alpha_j = 1$.
- First best is then given as a solution to

$$\max_{\left\{[c_{j,t}, l_{j,t}]_{j=1}^{N}\right\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\sum_{j=1}^{N} \alpha_j u_j(c_{j,t}, l_{j,t})\right]$$

subject to the resource constraint

$$\sum_{j=1}^{N} c_{j,t} \leq \sum_{j=1}^{N} l_{j,t} \text{ for all } t.$$

First Best (continued)

- Standard arguments imply that the *first-best allocation* satisfies

no distortions: $$\frac{\partial u_j(c_{j,t}^{fb}, l_{j,t}^{fb})}{\partial c} = -\frac{\partial u_j(c_{j,t}^{fb}, l_{j,t}^{fb})}{\partial l} \text{ for } j = 1, ..., N \text{ and all } t,$$

perfect smoothing: $c_{j,t}^{fb} = c_j^{fb}$ and $l_{j,t}^{fb} = l_j^{fb}$ for $j = 1, ..., N$ and all t.

Political Economy

- Let us model political economy with the following game form:

1. In each period t, we start with one party, j', in power.
2. All parties simultaneously make their labor supply decisions $l_{j,t}$. Output $Y_t = \sum_{j=1}^{N} l_{j,t}$ is produced.
3. Party j' chooses consumption allocations $c_{j,t}$ for each party subject to the feasibility constraint

$$\sum_{j=1}^{N} c_{j,t} \leq \sum_{j=1}^{N} l_{j,t}. \tag{11}$$

4. A first-order Markov process m determines who will be in power in the next period. The probability of party j being in power following party j' is $m(j \mid j')$, with $\sum_{j=1}^{N} m(j \mid j') = 1$ for all $j' \in \mathcal{N}$.

Political Economy (continued)

- Here MPE are straightforward and uninteresting (like Olson's roving bandits): maximum extraction every period by that group in power from all other groups.
- But Subgame Perfect Equilibria (SPE) potentially more interesting.
- These will have the feature that the current powerholder can be punished by high taxes/extraction in the future if it does not follow the agreed policy.
- The set of SPE is generally large, but one might wish to focus on the constrained Pareto efficient SPE (i.e., the frontier).
- Note that it is not possible to focus on a single point on this frontier, because as the identity of the group in power is stochastic, we will naturally move along this frontier.

Reevaluating Olson

- One possible question concerns the conditions under which first-best allocations are sustainable as SPE, and in particular how these depend on the Markov process for power switches.

Proposition: Consider an economy consisting of N groups. Suppose that $m(j \mid j) = \rho$ and $m(j' \mid j) = (1-\rho)/(N-1)$ for any $j' \neq j$. Then the set of sustainable first-best allocations is decreasing in ρ.

- Anti-Olson results. Why?
- Intuition: if $\rho = 0$ or very low, the group in power can be punished very strongly for extracting more than they are supposed to—next period they will not be in power with high probability, and they can be taxed very heavily.
- Conversely, when $\rho = 1$ or very high, less effective punishments, thus first-best publications are more difficult to sustain.

Local State Presence

- State capacity/presence is as much about local state presence.
- But in many countries there is endemic absence of the local and central state, such as Colombia.
- Rufino Gutierrez in 1912:

> *"...in most municipalities there was no city council, mayor, district judge, tax collector... even less for road-building boards, nor whom to count on for the collection and distribution of rents, nor who may dare collect the property tax or any other contribution to the politically connected..."*

- But then, local state presence doesn't just have direct effects. Indirect effects (spillovers) may be as important:
  - public good provision, policing, and law enforcement will impact neighboring municipalities.
  - possibility of strategic interactions, free riding or complementarities in investments.

Local Spillovers and Networks

- This suggests
  - Game theoretic model to understand interactions among municipalities and between municipalities and the national state in state capacity choices.
  - Estimate this model using data from Colombian municipalities to uncover:
    1. the **own effect** of state capacity on public goods and prosperity;
    2. the **spillover effects** of state capacity;
    3. the (strategic) **interaction effects** in state capacity choices (in particular, whether these are strategic complements or substitutes);
    4. the relationship between local and national state capacity choices.

Conceptual and Empirical Challenges

- Such an approach is to confront several key challenges:
  1. State capacity choices are endogenous.
  2. The estimation of spillovers ("contextual effects") is fraught with econometric difficulties because of reasons relate to both endogeneity and correlated effects.
  3. The estimation of strategic interactions is even more difficult, taking us to the territory of Manski's "endogenous effects".
- We will now discuss one attempt to overcome these problems in Acemoglu, Garcia-Jimeno and Robinson (2012).
- Strategy: structurally estimate model parameters using exogenous (historical) sources of variation.

Model of State Capacity over a Network

- Network game of public goods provision
- Interpret the administrative municipality-level map as a network:
  - Each municipality is a node
  - Each adjacency implies a link (undirected).
- Municipalities (and the national level) choose their levels of state capacity simultaneously.
- Utility functions are "reduced form" for a political economy process where state capacity has both costs and benefits.
- The national state has heterogeneous preferences over outcomes of different municipalities (cares more about some).

Model: Network Structure

- Let us represent the network with matrix $\mathbf{N}(\delta)$ with entries n_{ij} where

$$n_{ij} = \begin{cases} 0 & \text{if } j \notin N(i) \\ f_{ij} & \text{if } j \in N(i) \end{cases}$$

where

$$f_{ij} = \frac{1}{1 + \delta_1 d_{ij}(1 + \delta_2 e_{ij})}.$$

- $N(i)$ is the set of neighbors of i, d_{ij} is geodesic distance between i and j, e_{ij} is variability in altitude along the geodesic.
- In our benchmark, a network link is given by adjacency between municipalities.

Model: Technology

- We allow different dimensions of prosperity $j = 1, ...J$ to depend upon own and neighbors' state capacity:

$$p_i^j = (\kappa_i + \xi_i)s_i + \psi_1 s_i \mathbf{N}_i(\delta)\mathbf{s} + \psi_2^j \mathbf{N}_i(\delta)\mathbf{s} + \epsilon_i^j.$$

  where $\mathbf{N}_i(\delta)$ denotes the ith row of the network matrix.
- $\kappa_i + \xi_i$ is the effect of own state on own prosperity (heterogeneous, has observable and unobservable components);
- ψ_1 is the interaction effect (its sign determines whether this is a game of strategic complements or substitutes)
- ψ_2^j is a pure spillover effect from neighbors;

Model: State Capacity

- Let "state capacity" be a CES composite of locally chosen l_i and nationally decided b_i measures of state presence:

$$s_i = \left[\alpha l_i^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(\tau b_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \quad \sigma > 0, \tau > 0.$$

- Simplest is the special case where $\alpha = 0$. But we can also study the general case where $\alpha > 0$ and national bureaucracy also matters and is endogenously determined.

Model: Preferences

- Municipality i maximizes

$$U_i = \mathbb{E}_\epsilon \left[\frac{1}{J} \sum_j p_i^j - \frac{\theta}{2} l_i^2 \right].$$

- The national state maximizes

$$W_i = \mathbb{E}_\epsilon \left[\sum_i \left\{ U_i \zeta_i - \frac{\eta}{2} b_i^2 \right\} \right]$$

where the ζ_i are unobserved weights the national state puts on each i.

Model: Game

- National and local-level state capacities are chosen simultaneously. This gives us straightforward first-order conditions.
- Municipality choices:

$$\alpha \left[\frac{s_i}{l_i}\right]^{\frac{1}{\sigma}} \left[(\kappa_i + \xi_i) + \psi_1 \mathbf{N}_i(\delta)\mathbf{s}\right] - \theta l_i \begin{cases} < 0 & l_i = 0 \\ = 0 & , l_i > 0 \end{cases}.$$

- A game of strategic complements or substitutes depending on ψ_1:

$$\frac{\partial l_i}{\partial \mathbf{N}_i(\delta)\mathbf{s}} > 0 \Longleftrightarrow \psi_1 > 0$$

- We also derive the national state's first-order conditions (not shown here):
  - The main difference is that the national state does take into account spillovers, and weights municipalities heterogeneously.

Model: Linear Special Case

- When $\alpha = 1$ the game described above simplifies considerably.
- National-level choices no longer relevant, and $s_i = l_i$.
- Best responses become linear in neighbors' choices

$$s_i = \frac{\psi_1}{\theta}\mathbf{N}_i(\delta)\mathbf{s} + \frac{\kappa_i}{\theta} + \tilde{\xi}_i. \tag{12}$$

Proposition (Bramoulle, Kranton, and D'Amours (2012)): If $|\lambda_{min}(\mathbf{N}(\delta))| < \left(\frac{\psi_1}{\theta}\right)^{-1}$ the game has a unique Nash equilibrium.

- Notice that the reduced-form coefficient $\frac{\psi_1}{\theta}$ is analogous to what Manski (1993) calls an "endogenous effect"
  - Thus empirical work here must deal with the "reflection problem". Particularly serious since $\tilde{\xi}_i$'s likely to be spatially correlated.
- All of these results extend to the case where $\alpha < 1$.

Model: Identification Problem

- Substitute best responses into the prosperity equation

$$p_i^j = \theta s_i^2 + \psi_2^j \mathbf{N}_i(\delta)\mathbf{s} + \epsilon_i^j. \tag{13}$$

- In equilibrium, the spillovers, κ_i, and the interaction effect, ψ_1, drop out, and cannot be identified by running a regression of outcomes.
- This is because state capacity choices are (endogenously) a function of κ_i and ψ_1.
  - In addition, a quadratic relationship.
- Another identification challenge: ϵ_i^j's also likely to be spatially correlated.

Model: Idea for Identification

- Parameter κ_i a function of historical variables (described below) which are plausibly exogenous to the current development of state capacity and current prosperity.
  - Also, conveniently, they happen to be spatially uncorrelated.
- Using these variables and the network structure, estimate (12) and (13)—using linear IV, system GMM, maximum likelihood or simulated method of moments (SMM).
- From (12), we estimate $\frac{\psi_1}{\theta}$ (from the endogenous effect) and (local) average κ_i (from the intercept).
- From estimate (13), we estimate θ and ψ_2, fully identifying all of the parameters.
- Note the importance of estimating the endogenous effect, from which the crucial identification of the outcome equation comes.

Empirical Strategy: Traditional Approaches

- This is a network identification problem.
- Commonin the literature are two strategies. First, assume away correlated effects.
- Second, exploit network structure to break the "reflection problem".
- Most creatively: Bramoulle et al (2009):
  - If for every node i $\exists k$ such that $k \in N(j)$, $j \in N(i)$, and $k \notin N(i)$, then covariate x_k is a valid instrument.
  - Thus use powers of $\mathbf{N}_i(\delta)\mathbf{x}$ as instruments for s_i.
  - But problem: we may not know network structure exactly and more importantly, correlated effects that extend beyond immediate neighborhood.

Empirical Strategy: Exclusion Restrictions

- Better identification strategy would be to exploit exogenous sources of variation both for own and spillover effects.
- Idea: use historical variables affecting the development of the state: colonial stage presence and royal roads, denoted by vector $\mathbf{c}$.
- Formally:

$$cov(\mathbf{N}_i(\delta)\mathbf{c}, \tilde{\xi}_i) = cov(\mathbf{N}_i^2(\delta)\mathbf{c}, \tilde{\xi}_i) = 0$$

and

$$cov(\mathbf{c}, \epsilon_i^j) = cov(\mathbf{N}_i(\delta)\mathbf{c}, \epsilon_i^j) = cov(\mathbf{N}_i^2(\delta)\mathbf{c}, \epsilon_i^j) = 0.$$

- Why is this plausible?

Colonial State Presence

- Highly concentrated colonial state presence around key cities and resources:
  - Colonial state presence in gold mining regions related to taxation purposes.
  - Colonial state presence in high native population regions related to control of the population, legal adjudication, etc.
  - Colonial state presence in geographically strategic places related to military aims.
- Gold mining, native populations and those military aims are no longer relevant. So the direct effect of colonial state presence is by creating the infrastructure for current state presence.

Royal Roads

- Royal roads were one of the few investments in infrastructure (building upon pre-colonial roads).
- The presence of royal roads is a good indicator of where the colonial state was interested in reaching out, and controlling territory.
- But most of these royal roads were subsequently abandoned as transportation infrastructure.
  - Most of these were built for porterage along difficult geography, making them hard to subsequently reconvert to new transportation technologies.
- Good case that these are excludable (especially conditional on current road network).

Correlated Effects

- If these instruments are spatially correlated, then the spatial correlation of current outcomes might project on them, leading to bias.
- Interestingly, very little spatial correlation of the colonial state presence or royal roads data.

Table 2. Within-department spatial correlation of historical state presence variables

	1	2	3	4	5	6	7	8	9
1. Own distance to royal roads	1.000								
2. Neighbors' average distance to royal roads	0.283	1.000							
3. Neighbors of neighbors' distance to royal roads	0.045	0.615	1.000						
4. Own colonial officials	-0.095	-0.072	-0.047	1.000					
5. Neighbors' average colonial officials	-0.146	0.039	0.060	-0.061	1.000				
6. Neighbors of neighbors' colonial officials	-0.044	0.063	0.072	-0.062	-0.070	1.000			
7. Own colonial state agencies	-0.135	-0.039	-0.017	0.545	-0.006	-0.002	1.000		
8. Neighbors' average colonial state agencies	-0.208	0.250	0.283	-0.053	0.490	0.008	0.022	1.000	
9. Neighbors of neighbors' colonial state agencies	-0.193	0.244	0.334	-0.036	0.031	0.408	0.078	0.289	1.000

Correlations reported are the average across-departments of the correlations for each department.

Empirical Strategy: Linear Model

- Now focus on $\alpha = 1$ (linear best responses and prosperity equations).
- Suppose that $\frac{\kappa_i}{\theta} = g(\mathbf{c}_i\boldsymbol{\varphi} + \mathbf{x}_i\boldsymbol{\beta}) + \varsigma^D$, where ς^D are department fixed effects.
- Then we have

$$s_i = \frac{\psi_1}{\theta}\mathbf{N}_i(\delta)\mathbf{s} + g(\mathbf{c}_i\boldsymbol{\varphi} + \mathbf{x}_i\boldsymbol{\beta}) + \varsigma^D + \tilde{\xi}_i.$$

$$p_i^j = \theta s_i^2 + \psi_2^j \mathbf{N}_i(\delta)\mathbf{s} + \mathbf{x}_i\tilde{\boldsymbol{\beta}}^j + \tilde{\varsigma}^{jD} + \epsilon_i^j.$$

- Two econometric strategies
  - Linear IV (normalizing $\delta = (1, 1)$) estimate these equations separately.
  - System GMM ($J + 1$ equations) that exploits the joint dependence on θ, and allows for estimation of δ.

Results: First Stage for Best Response

Panel II		First stage for $N_i(\delta)$s							
Neighbors' colonial state officials		0.320	0.338			0.556	0.637		
		(0.096)	(0.100)			(0.143)	(0.155)		
Neighbors' colonial state agencies		1.275	1.242			1.673	1.631		
		(0.126)	(0.131)			(0.211)	(0.223)		
Neighbors' distance to royal road		-1.031	-0.992			-1.497	-1.456		
		(0.219)	(0.223)			(0.278)	(0.287)		
Neighbors of neighbors' colonial state officials		0.209	0.269			0.311	0.427		
		(0.170)	(0.177)			(0.240)	(0.258)		
Neighbors of neighbors' colonial state agencies		0.649	0.568			1.085	0.937		
		(0.181)	(0.190)			(0.264)	(0.281)		
Neighbors of neighbors' distance to royal road		0.178	0.172			0.268	0.296		
		(0.169)	(0.173)			(0.231)	(0.236)		
First-stage R-squared:		0.681	0.671			0.681	0.658		
F-test for excluded instruments:		17.0	145.6			19.55	171.0		
F-test p-value		0.000	0.000			0.000	0.000		
Overidentification test:	Test statistic	4.053	6.350			4.399	5.775		
Chi-squared(2)	P-value	0.542	0.385			0.494	0.449		
Log population		Control	Control	Instrum	Instrum	Control	Control	Instrum	Instrum
Observations		975	975	975	963	1017	1017	1017	1003

Interpretation

- Very strong first stage.
- Overidentification tests never reject the validity of subsets of instruments.
- Effects plausible:
  - Neighbors' colonial state officials and agencies significantly increase neighbors' state capacity today.
  - Neighbors' distance to royal roads significantly reduce neighbors' capacity today.
  - Weaker but still significant effects of neighbors of neighbors.

Estimates of the Best Response Equation

Table 3. Contemporary State Equilibrium Best Response

State capacity measured as log of:	Number of state agencies				Number of municipality employees			
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	IV	Sys. GMM	OLS	IV	IV	Sys. GMM
	Equilibrium best response							
ds_i/ds_j	0.016	0.017	0.019	0.020	0.021	0.022	0.022	0.016
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
ds_i/dcolonial state $officials_i$	0.127	0.128	0.108	-0.040	0.129	0.130	0.105	0.087
	(0.031)	(0.031)	(0.033)	(0.050)	(0.043)	(0.043)	(0.047)	(0.069)
ds_i/dcolonial state $agencies_i$	0.003	0.001	-0.016	0.096	0.017	0.017	-0.002	0.085
	(0.033)	(0.033)	(0.033)	(0.055)	(0.058)	(0.059)	(0.060)	(0.085)
ds_i/ddistance to royal $road_i$	0.008	0.010	0.007	0.074	-0.035	-0.035	-0.038	-0.036
	(0.019)	(0.019)	(0.021)	(0.034)	(0.034)	(0.035)	(0.036)	(0.044)

Interpretation

- Best response slopes upward (investments in state capacity are strategic complements).
- Own colonial state officials significantly increase own state capacity today.
- Conditional on this, colonial state agencies and distance to royal roads insignificant, but significant with the right sign when colonial state officials are excluded.

Estimates of the Prosperity Equation

Table 4A. Prosperity and Public Goods Structural Equation

State capacity measured as: log of number of municipality state agencies

Dependent variable	Life quality index				Public utilities coverage			
Panel I	(1) OLS	(2) IV	(3) IV	(4) Sys. GMM	(5) OLS	(6) IV	(7) IV	(8) Sys. GMM
	Prosperity equation							
dp_i/ds_i	0.802	0.394	0.389	0.314	0.602	0.563	0.567	0.314
	(0.044)	(0.135)	(0.143)	(0.041)	(0.037)	(0.127)	(0.134)	(0.041)
dp_i/ds_j	0.015	0.024	0.025	0.025	0.022	0.020	0.020	0.027
	(0.004)	(0.006)	(0.006)	(0.004)	(0.004)	(0.006)	(0.006)	(0.003)
Panel II	First stage for s_i^2							
F-test for excluded instruments:		31.23	35.39			31.01	35.06	
F-test p-value		0.000	0.000			0.000	0.000	
First-stage R-squared		0.670	0.655			0.670	0.655	
First-stage linear model:	First stage for $N_i(\delta)$s							
F-test for excluded instruments:		526.7	523.7			524.6	522.1	
F-test p-value		0.000	0.000			0.000	0.000	
First-stage R-squared		0.769	0.770			0.769	0.770	
Log population	Control	Control	Instrum	Instrum	Control	Control	Instrum	Instrum
Observations	973	973	973	963	975	975	975	963

Estimates of the Prosperity Equation (continued)

State capacity measured as: log of number of municipality state agencies

Dependent variable	Not in poverty				Secondary enrollment			
Panel I	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	OLS	IV	IV	Sys. GMM	OLS	IV	IV	Sys. GMM
	Prosperity equation							
dp_i/ds_i	0.520	0.342	0.353	0.314	0.515	0.178	0.223	0.314
	(0.038)	(0.141)	(0.147)	(0.041)	(0.049)	(0.179)	(0.186)	(0.041)
dp_i/ds_j	0.019	0.021	0.021	0.021	0.023	0.036	0.035	0.035
	(0.004)	(0.006)	(0.006)	(0.003)	(0.005)	(0.007)	(0.007)	(0.004)
Panel II	First stage for s_i^2							
F-test for excluded instruments:		31.01	35.06			30.46	35.70	
F-test p-value		0.000	0.000			0.000	0.000	
First-stage R-squared		0.670	0.655			0.675	0.662	
First-stage linear model:	First stage for $N_i(\delta)s$							
F-test for excluded instruments:		524.6	522.1			579.3	583.1	
F-test p-value		0.000	0.000			0.000	0.000	
First-stage R-squared		0.769	0.770			0.771	0.773	
Log population	Control	Control	Instrum	Instrum	Control	Control	Instrum	Instrum
Observations	975	975	975	963	965	965	965	963

Interpretation and Magnitudes

- Own effect more than 10 times the impact on neighbors, which is plausible.
- But the external it is on several neighbors, so the partial equilibrium spillover and direct effects comparable.
- But full equilibrium effects, factoring in endogenous responses, indicate much larger network effects than direct effect.

Evaluating Quantitative Magnitudes

Table 5. Experiment: Implications of Moving All Municipalities below Median State Capacity to Median

Panel Ia	Linear model									
Partial equilibrium change in:	Local agencies:		Life quality index		Utilities coverage		% not in poverty		Secondary enroll.	
	From	To	From	To	From	To	From	To	From	To
Change in median:	10	10	48.0	49.0	53.3	57.2	57.1	60.0	56.6	59.2
Fraction due to own effect:			53.4%		51.7%		57.1%		45.5%	
Fraction due to spillovers:			46.6%		48.3%		43.0%		54.5%	

Panel Ib										
General equilibrium change in:	Local agencies:		Life quality index		Utilities coverage		% not in poverty		Secondary enroll.	
	From	To	From	To	From	To	From	To	From	To
Change in median:	10	20.6	48.0	58.2	53.3	73.7	57.1	68.3	56.6	82.4
Fraction due to direct effect:			9.8%		18.9%		25.5%		10.1%	
Fraction due to network effects:			90.2%		81.1%		74.5%		89.9%	

Evaluating Quantitative Magnitudes (continued)

- This implies that increasing local state presence in all municipalities below the median to the median of the country, holding all other state capacity choices fixed, reduces fraction of the population about poverty from 57% to 60%.
- About 57% of this increase is due to direct effects, and the remaining 43% to spillovers.
- But once, there is this induced change in state capacities, there will be further network responses—through strategic complementarities.
- Once these are factored in, fraction of the population above poverty rises to 68%— of course all of this due to network effects.

Robustness

- Very similar results with System GMM.
- The $g(\cdot)$ function is indeed nonlinear, but implied quantitative magnitudes are essentially the same.
- Similar results with additional controls and to different strategies, including various controls for the effects of national bureaucracy.

Monopoly of Violence

- Max Weber defined a state as "a human community that (successfully) claims the *monopoly of the legitimate use of physical force* within a given territory."
- Many states do not have such a monopoly and without it have little hope of enforcing rules, regulations, and laws, providing property rights and public goods.
- Presumption in the existing literature: this is because of the weakness of the state and 'modernization' will ultimately strengthen the state and ensure monopoly of violence.

But

- In many polities, the central state exists side-by-side, and in fact in a '*symbiotic*' relationship with non-state armed actors.
- Examples:
  - Waziristan in Pakistan;
  - Kurdish areas in Iraq;
  - the Mafia in the south of Italy;
  - Southern United States after the Hayes-Tilden agreement of 1877;
  - Colombia.
- Weak states or *symbiotic* relationship between central state and periphery?

Weak vs. Strong States in Colombia

- Imagine there is an incumbent politician/party facing an election.
- The country is divided into regions some of which are controlled by non-state armed actors.
- The incumbent decides which regions to 'take back' (in the limit establishing a monopoly of violence) and chooses a policy vector in the election.
- Non-state armed actors have preferences over policies and can coerce voters to support one candidate over another.
- This creates an electoral advantage for incumbent politicians they favor and reduces the incentives to eliminate these non-state actors.

Formal Model

- Let us consider a two-period model of political competition between two parties.
- Party A is initially (at $t = 0$) in power and at $t = 1$, it competes in an election against party B.
- The country consists of a large equal-sized number, N, of regions, with each region inhabited by a large number of individuals. We denote the collection of these regions by $\mathcal{N}$.
- The party that wins the majority of the votes over all regions wins the election at the time $t = 1$.
- Regions differ in terms of their policy and ideological preferences and, in addition, some regions are under paramilitary control.
- We assume as in standard Downsian models that parties can make commitments to their policies, but their ideological stance is fixed and captures dimensions of policies to which they cannot make commitments.

Preferences

- The utility of individual i in region $j \in \mathcal{N}$ (i.e. $j = 1, ..., N$) when party $g \in \{A, B\}$ is in power is given by

$$U_{ij}\left(q, \tilde{\theta}^g\right) = u_j\left(q\right) - Y\left(\tilde{\theta}_j - \tilde{\theta}^g\right) + \tilde{\varepsilon}_{ij}^g,$$

  where $q \in Q \subset \mathbb{R}^K$ is a vector of policies, u_j denotes the utility of individuals in region j, $\tilde{\theta}_j$ is the ideological bliss point of the individuals in region $j \in \mathcal{N}$, so that $Y\left(\tilde{\theta}_j - \tilde{\theta}^g\right)$ is a penalty term for the ideological distance of the party in power and the individual.
- Finally, $\tilde{\varepsilon}_{ij}^g$ is an individual-specific utility term where

$$\tilde{\varepsilon}_{ij}^A - \tilde{\varepsilon}_{ij}^B = \xi + \varepsilon_{ij},$$

  where ξ is a common valance term and ε_{ij} is an iid term.
- ξ and each ε_{ij} have uniform distributions over $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$.

Summary of Results

- **Benchmark model without paramilitaries:** standard symmetric equilibrium with policy convergence to the average preference of N regions, but the party that is ideologically more popular among voters wins with higher probability.
- **Equilibrium with "passive" paramilitaries:** still policy convergence but now policies targeted to voters in non-paramilitary areas and the party that is ideologically more popular among the paramilitaries (in addition to being more popular among voters in non-paramilitary areas) wins with higher probability.

Summary of Results (Continued)

- **Persistence of paramilitaries:** now suppose that the state can reconquer areas under paramilitary control; whether it would like to do so will depend, in part, on whether the paramilitaries support the party currently controlling the government.
  - *Most importantly*: if the current government is popular with paramilitaries, it will be less likely to reconquer areas where paramilitaries are likely to provide the decisive votes for its victory.
- **Equilibrium with "active"' paramilitaries:** now paramilitaries endogenously decide which party to support; similar results, except that now both parties will change their platforms to "appease" paramilitaries.

Implications

- *Main implication:* paramilitaries will tend to persist to the extent that they deliver votes to the incumbent executive and that this effect is larger in areas where the President would otherwise not do well.
- Thus non-state armed actors can persist because they can be in a *symbiotic relationship* with the executive.
  - On the one hand, paramilitaries deliver votes to the President and in addition elect legislators who support the executive.
  - On the other, the executive delivers laws and the policies that the paramilitaries prefer.
- In addition, policies chosen to appease paramilitaries rather than provide public goods and services to the population.

Some Colombian Background

- In recent years Colombia has been dominated by two main non-state armed actors:
  - the 'left-wing' Fuerzas Armadas Revolucionarias de Colombia (FARC—The Revolutionary Armed Forces of Colombia) and
  - the 'right-wing' paramilitary forces which in 1997 coalesced into the Autodefensas Unidas de Colombia (AUC—United Self-Defense Organization of Colombia).
- After the foundation of the AUC in 1997 a strategic decision was taken to influence national politics (possibly taken at Santa Fé de Ralito in 2001 where members of the AUC, politicians and members of congress signed a document calling for the 'refounding of the country.')

The Involvement of Paramilitaries in Politics

- In 2005 accusations of involvement of the AUC in the elections of 2002. Scandal with the demobilization of Jorge 40 and his 2,000 strong block on March 10, 2006 in La Mesa, César.
- Jorge 40's computer fell into the hands of government officials and it contained emails ordering his men to recruit peasants to pretend to be paramilitaries during demobilization ceremonies and also listed over 500 murders, and many links between politicians and paramilitaries.
- So far around 30,000 paramilitaries have "demobilized" in this process.
- As of April 22, 2008, 62 members of Congress and the Senate were official suspects, 33 lawmakers, including Mario Uribe, President Uribe's cousin, were in jail awaiting trial for links with paramilitaries.

Who Are the Representatives?

Table 1: Top 20 Senators by Vote Share in Paramilitary Presence Areas

Senator	Third Parties	Votes Param. Zones	Reelection	Just.& Peace Law	Investigated/Arrested
GNECCO	Yes	61.88	Yes	Yes	No/No
PIMIENTO	Yes	54.6	Yes	Yes	No/Yes
MALOOF	Yes	49.5	Yes	Yes	No/Yes
CLAVIJO	Yes	44.76			No/Yes
SAADE	Yes	42.51	Yes		Yes/No
MARTINEZ	Yes	41.8	Yes		Yes/No
GARCIA	Yes	38.01			No/Yes
PUELLO	No	30.64	Yes		No/No
MERLANO	Yes	28.73	Yes	Yes	No/Yes
VIVES	Yes	27.52	Yes	Yes	No/Yes
MONTES	Yes	26.9	Yes	Yes	No/No
ZUCCARDI	Yes	25.09	Yes	Yes	No/No
ARAUJO	Yes	24.44		Yes	No/Yes
SERRANO	No	23.12	No		No/No
QUINTERO	Yes	23.06			No/Yes
DE LA ESPRIELLA	Yes	22.55	Yes	Yes	No/Yes
ACOSTA	Yes	22.4			No/No
BLEL	Yes	21.59	Yes	Yes	No/Yes
GIL	Yes	21.21	Yes		No/Yes
CORDOBA	No	20.14	No	No	No/No

Impact of Paramilitaries on Elections

- Acemoglu, Robinson and Santos (2013) show that after the decision of the AUC to enter politics (i.e., after 2001), the presence of paramilitaries in a municipality is robustly correlated with greater vote shares of 'third parties,' typically connected with paramilitaries and supporting right-wing positions.
- Also paramilitary presence correlated with 'electoral concentration' for Senate and Congress elections in 2002 and the vote share of Presidential Uribe, in 2002 and more so in 2006.
- No robust effects of guerilla presence on voting patterns.
- Generally, the data support the idea that paramilitaries have a large impact on elections. Consistent with the case study literature.

Regression Evidence

TABLE 3. Paramilitary presence and third parties share of votes in the elections for the senate.

Dependent Variable is Vote Share Obtained by Third Parties in the Elections for the Senate	Panel 1991-2006 (1)	Panel 1991-2006 (2)	Panel 1991-2006 (3)	Panel 1991-2006 (4)	Panel 1991-2006 (5)	Panel 1991-2006 (6)
	Armed Actors Presence is Measured by:					
	Attacks			Attacks Dummy		
Paramilitary Presence						
Paramilitary Presence X 1994	10.04	3.13	0.05	5.26	2.42	1.93
	(2.956)	(2.899)	(2.956)	(1.137)	(1.127)	(1.189)
Paramilitary Presence X 1998	12.54	5.20	2.98	5.67	2.49	2.86
	(4.118)	(4.653)	(4.719)	(1.508)	(1.543)	(1.615)
Paramilitary Presence X 2002	42.22	31.85	27.82	16.59	11.73	11.12
	(5.944)	(5.639)	(5.764)	(1.815)	(1.798)	(1.849)
Paramilitary Presence X 2006	38.66	20.89	16.86	16.18	7.99	7.52
	(5.478)	(5.090)	(5.063)	(1.856)	(1.604)	(1.666)
Guerrilla Presence						
Guerrilla presence X 1994			1.73			1.70
			(0.939)			(1.458)
Guerrilla Presence X 1998			1.25			−1.28
			(0.993)			(1.777)
Guerrilla Presence X 2002			2.27			2.16
			(1.264)			(1.940)
Guerrilla Presence X 2006			2.26			1.60
			(1.091)			(1.784)
Controls Interacted with Year Dummies	No	Yes	Yes	No	Yes	Yes
Observations	5379	4915	4915	5379	4915	4915

Regression Evidence (continued)

TABLE 4. Paramilitary presence and winning presidential candidate share of votes.

Dependent Variable is Winning Presidential Candidate Vote Share	Panel 1998-2006 (1)	Panel 1998-2006 (2)	Panel 1998-2006 (3)	Panel 1998-2006 (4)	Panel 1998-2006 (5)	Panel 1998-2006 (6)
	Armed Actors Presence is Measured by:					
	Attacks			Attacks Dummy		
Paramilitary Presence						
Paramilitary Presence X 2002	13.75	6.71	10.91	4.55	1.26	2.23
	(4.310)	(2.720)	(2.717)	(1.393)	(1.112)	(1.159)
Paramilitary Presence X 2006	37.19	24.71	17.85	10.86	5.92	3.65
	(5.584)	(3.283)	(3.307)	(1.568)	(1.197)	(1.201)
Guerrilla Presence						
Guerrilla Presence X 2002			−2.39			−3.43
			(0.634)			(1.121)
Guerrilla Presence X 2006			3.86			7.85
			(0.927)			(1.390)
Controls Interacted with Year Dummies	No	Yes	Yes	No	Yes	Yes
Observations	3297	2951	2951	3297	2951	2951

Regression Evidence (continued)

	Cross Section (1)	Cross Section (2)	Cross Section (3)	Cross Section (4)	Cross Section (5)	Cross Section (6)	Cross Section (7)	Cross Section (8)
	Dependent Variable is the Fraction of Arrested Senators/Congressmen in list l.							
	Senate				Congress			
Dummy Conservative	0.05			0.09	−0.01			0.09
	(0.117)			(0.139)	(0.06)			(0.09)
Dummy Left	−0.10			−0.03	−0.06			−0.02
	(0.058)			(0.073)	(0.03)			(0.04)
Dummy Third Parties	0.21			0.22	0.06			0.08
	(0.089)			(0.085)	(0.05)			(0.05)
Share of Votes From:								
Paramilitary Areas		1.43	1.13	1.01		0.26	0.20	0.18
		(0.413)	(0.493)	(0.431)		(0.12)	(0.13)	(0.13)
Guerrilla Areas		−0.24	0.15	0.30		−0.01	0.03	0.07
		(0.727)	(0.803)	(0.740)		(0.07)	(0.07)	(0.08)
Right Oriented Areas			−0.33	−0.41			−0.31	−0.44
			(0.415)	(0.373)			(0.12)	(0.21)
Left Oriented Areas			−0.24	−0.27			−0.09	−0.10
			(0.183)	(0.183)			(0.06)	(0.06)
Observations	96	96	96	96	162	162	162	162
R-squared	0.072	0.159	0.174	0.241	0.01	0.04	0.06	0.08

Further Evidence

- Consider the vote in 2005 in the senate to re-introduce two clauses of the Justice and Peace Law which had been vetoed in Congress. These two clauses were to stop former paramilitaries being charged with sedition (avoiding possible extradition), and a limit of the length of prison services they could serve.
- Evidence that the presence of paramilitaries in areas where senate lists received a high proportion of their votes helps to predict the way Senators on the list vote.
- This variable also predicts which Senators were subsequently arrested for connections with paramilitaries.
- See Table 1.

Main Result: Persistence of the Paramilitaries

- Consider now whether or not the persistence of paramilitaries after the 2002 election is related to voting patterns in 2002.
- In line with the predictions discussed above, paramilitaries tend to persist more in a municipality, the greater the vote share of President Uribe in 2002. This effect is smaller, the greater was the historical extent of conservative support in the municipality.
- The intuition for this last finding is that in places with strong historical support for conservatives Uribe was confident of winning and therefore needs the support of paramilitaries less.

Regression Evidence

TABLE 6. Persistence of paramilitaries and vote share for Alvaro Uribe.

Dependent Variable is Paramilitary Attacks in 2003-2005	Cross-Section (1)	Cross-Section (2)	Cross-Section (3)	Cross-Section (4)	Cross-Section (5)	Cross-Section (6)
	Sample is Restricted to Municiaplities with Paramilitary Presence in 1999-2001					
Uribe Vote Share	0.14	0.13	0.13	0.17	0.15	0.14
	(0.084)	(0.096)	(0.097)	(0.088)	(0.099)	(0.097)
Uribe Vote Share X Pastrana Vote Share	−0.77	−0.62	−0.63	−1.53	−1.36	−1.42
	(0.357)	(0.386)	(0.410)	(0.545)	(0.619)	(0.662)
Patrana Vote Share	−0.23	−0.07	−0.08	−0.39	−0.04	−0.05
	(0.087)	(0.123)	(0.125)	(0.144)	(0.166)	(0.166)
Paramilitary Presence in 1999-2001	0.41	0.40	0.40	0.34	0.34	0.35
	(0.160)	(0.170)	(0.187)	(0.152)	(0.160)	(0.178)
Guerrilla Presence in 1999-2001			−0.00			−0.01
			(0.025)			(0.025)
Controls	No	Yes	Yes	No	Yes	Yes
Quartic on Pastrana's Vote Share	No	No	No	Yes	Yes	Yes
F test on Pastrana's Vote Share Higher Order Terms				4.66	2.73	2.83
P Value of F test				0.003	0.044	0.038
Observations	319	309	309	319	309	309
R-squared	0.21	0.24	0.24	0.259	0.274	0.276

The Symbiotic Relationship

- Large case study literature suggests that Uribe has delivered to paramilitaries (lenient Justice and Peace Law, refusal to support suspension of alternates in the legislature for arrested politicians).
- One can look at the roll call vote in 2004 to change the constitution to allow for Presidential re-election.
- Result: the greater the proportion of votes a Senate list received in paramilitary areas, the greater the proportion of Senators on the list that voted for re-election.

Persistence of Civil Wars

- Related idea in Acemoglu, Ticchi and Vindigni: civil wars, and thus the lack of power of the central state in many parts of the territory, persist in many African countries because the elites are afraid of increasing the power of the military.
  - Extreme example: Sierra Leone.
- The elite (often the politicians controlling the coffers) have a choice:
  - Increase the power of the military. This will end the civil war, but then will unleash a series of further political changes, either necessitating greater power sharing with the military, or other political reforms.
  - Keep the military weak. This will lead to the persistence of civil wars, but the elite can still grab a little rents.