Political Accountability in Divided Societies: The Politics of Fear

Gerard Padró i Miquel*
Massachusetts Institute of Technology
Department of Economics

October, 2004
JOB MARKET PAPER

Abstract

I present a model in which the combination of divided societies and weak institutions in the form of "Personal Rule" regimes creates a failure of accountability of the leadership. As a consequence, the ruler can construct a kleptocratic regime in which he can steal from the citizens, included his own supporters, even though he needs them to survive politically. The model also predicts extensive use of patronage and absence of investment in infrastructure. Hence, the model fits the experience of bad governance, wasteful policies and kleptocracy in post-colonial Africa.

*I wish to thank Daron Acemoglu, Abhijit Banerjee and Jim Snyder for encouragement and advice. I thank George-Marios Angeletos, Robert Bates, Kanchan Chandra, Drew Fudenberg, Luis Garicano and Iván Werning for helpful conversations. For suggestions and comments I am grateful to Pol Antràs, Raphael Auer and several participants in the Harvard/MIT Development Seminar and the NBER Summer Institute. All remaining errors are mine. I gratefully acknowledge financial support from the Fundación Ramón Areces.
1 Introduction

“Very quickly, African independence and socialism turned into one-man dictatorships, characterized by conspicuous consumption by the elites and a "Swiss-bank socialism" that allowed the head of state and his cohort of vampire elites to loot their countries’ treasuries (...). Billions of dollars were deposited abroad by the Babangidas, the Bandas, the Barres, the Does, the Kerekous, the Houphouet-Boignies, the Mois, the Mobutus, the Mengistus and many others.”

The plundering of African economies has been a systematic and generalized practice of African leaders in the years since their countries achieved independence. The private rewards from such activities have been enormous: several rulers, such as Mobutu, Moi or Houphouet-Boigny, have been estimated to have personal fortunes equivalent to the total external debt accumulated by their countries. Accompanying the blatant enrichment of the leadership, there has been an effort to engage in extensive redistribution of resources in surprisingly inefficient ways. For instance, Bates (1981) describes how cash crop growers are taxed by being forced to sell their produce to state-owned Marketing Boards at prices well below the international price to finance the inefficient industrial sector. These manipulations of the market extend to other policies such as financial repression, subsidized loans, import licences, restrictions in the holdings of foreign currency and even banning of interregional trade in foodstuffs. In addition to the implied efficiency costs, this array of regulatory practices has provided a reason for the creation of bloated and inefficient bureaucracies that consume a huge share of public resources and are used as distribution channels for patronage.

Extensive and inefficient redistribution and regulation, corruption and outright stealing by the leadership explain a large share of the dismal economic performance of African economies in the last five decades. Despite the exploitation of seemingly boundless natural resources, the reception of important amounts of foreign aid and an astronomical increase of foreign debt, most of Africa’s population earns today less income than the already low 1960 levels. Intuition suggests that a leader who reduces the utility of her citizens should find it very difficult to remain in power. The ability of the citizenry to replace non-performing politicians should constrain the leaders to dedicate their efforts,

\[1\] Quote from Ayittey (1992:105)
\[2\] See Ayittey (1992) or Mbaku (2000) for an account of the extend of thievery.
\[3\] For the effect of bad policies, see Collier and Gunning (1999), Easterly and Levine (1997) or Easterly (2002). For the effect of political and bureaucratic corruption see, for instance, Mbaku (2000).
to some extent, to the advancement of the interests of their constituencies. Nonetheless, the facts in Africa suggest that this mechanism fails to work: the political longevity of kleptocratic and inefficient rulers is remarkable. For instance, Kenya endured Moi’s leadership for 24 years, and Houphouet-Boigny’s leadership of Ivory Coast ended with his death, after 33 years in power. What explains the incapacity of the citizenry to replace or at least constrain such blatantly venal leaders?

A standard explanation is that coercion and force are used to crush any potential dissent. However, African states -colonial, and post-colonial alike- are generally portrayed as very weak states, with limited control over their territories, with weak armies and police forces and severely limited bureaucracies. This is difficult to reconcile with the idea that such regimes have been able to keep their populations from expressing their discontent at the enrichment of the few and the poverty of the many solely through the use of violence. In other words, African states have by no means been police states, with tight control over the private activities of citizens. It is precisely this weakness that explains the incidence and longevity of guerrilla movements and other forms of violent conflict. Moreover, the most successful kleptocratic regimes, such as Zaire or Kenya typically never resorted to widespread violence.

Instead, a better explanation is that these leaders have been able to co-opt large shares of their populations into supporting the regime. The patterns of inefficient redistribution can thus be rationalized by the need to obtain support from sizable groups. However, this explanation opens a new set of questions: how is it possible that the leaders were able to amass such exorbitant personal wealth at the same time that they were rewarding a share of their populations sufficiently large to maintain power? Even more intriguing, if the authority of the ruler is cemented by the support of a particular group, how is it possible that internal competition within the group has not dissipated, at least to some extent, the enormous rents that these leaders were accumulating?

This paper presents an argument to explain these policy patterns together with the

---

4 The seminal theoretical work on these lines is Barro (1973) and Ferejohn (1986). These models are based on an underlying moral hazard problem that allows the ruler some room to shirk in office. Optimal use of replacement rules via election of another candidate places limits on this potential shirking.


7 Note that an explanation based on oppression results in leaders insulated from the population and secure in their power. If this were the case, why would they engage in such extensive and inefficient transfers of resources across groups in their societies?

8 It is surprising to see many instances were large populations are easily mobilized by the regime: ethnic voting is such an example, but the much darker side of it was present in Rwanda, in 1994.
absence of political accountability and the endurance of kleptocratic regimes. The explanation relies on the characterization of African polities as “Personal Rule”\(^9\) regimes, a particular type of “weakly institutionalized”\(^10\) political systems.

Specifically, the model uses three consequences of weak institutionalization. First, checks and balances are basically nonexistent. This allows leaders to use their discretionary powers to treat different groups in their societies in discriminatory ways, effectively constructing clientelist networks. In Africa, the group attribute that is typically most politically relevant is the ethnic dimension\(^11\).

Ethnic groups feature two characteristics that are essential for the argument presented here. Firstly, as developed extensively in Bates (1983), in Africa ethnic groups are the result of competition for the spoils of modernization, and, as a consequence, they serve as exclusion devices. This means that it is very difficult for an individual to change his ethnic definition or allegiance, especially if it is to join the group of “winners.” This phenomenon creates intraethnic loyalty because in the short to medium term, the fate of each individual is linked to the fate of the group. This constrain is true both for the rank-and-file and the elite members of an ethnic group. Secondly, as a consequence of this last point, ethnic links provide a two-way commitment device for the future. The ruler can commit to provide patronage to his ethnic supporters in the future, and supporters can commit to support their leader\(^12\). Moreover, since the strength of personal links inside the group is stronger than across groups, deviations are easier to punish if the group is responsible for it\(^13\), making intra-group cooperation easier to sustain.

Second, in a system of personal rule the process of replacing a leader in power is not controlled by an established political institution, and hence it is uncertain and hardly ever smooth. Jackson and Rosberg (1982), write “a succession [...] alters at least some of the important relationships and standings among leaders and factions- for example, the standing of big men and the clan and ethnic communities they represent” and “The ultimate uncertainty in a system of personal rule lies in the key point of vulnerability: the ruler. [...] If he falls, his relatives, friends, lieutenants, clients and followers also may fall, and the ensuing political disruption may threaten the political peace.” Thus, ousting a leader may initiate a process that involves a potential change in the relative

---

\(^9\)See Jackson and Rosberg (1982) for a characterization of such regimes.


\(^11\)For a comment on the political relevance of ethnic categories see Horowitz (1985). Bates (1983) comments on the particularities of ethnicity in Africa. Note that Africa is the continent with the highest degree of ethnolinguistic fractionalization.

\(^12\)For other instances of ethnic groups as commitment devices see Bates (2000).

\(^13\)See Fearon and Laitin (1996) and Miguel and Gugerty (2004) for different arguments based on ethnic groups enhancing the capacity to punish deviations by group members.
status of different groups in society. The cases of Kenya or Cameroon are good examples of dramatic switches in the standing of different ethnic groups caused by succession.

Finally, a weak state is forced to use different forms of indirect taxation as its main mechanism for raising revenue\textsuperscript{14}.

Due to the absence of constraints, the ruler can easily create first and second class citizens, treating her ethnic group better than the opposition. The kleptocratic equilibrium is sustained because the way regime supporters constrain the rapacious activities of the leadership is by credibly threatening to replace the ruler and put someone else in charge, \textit{ideally from their own group}. However, retiring support from the ruler initiates an uncertain succession process that may end up with the current supporting group in the opposition. The ethnic divide implies that the division between winners and losers is very difficult to cross ex-post\textsuperscript{15} and hence the possibility of losing the winner status is not satisfactory: being in the opposition is necessarily worse than being in the supporter group, because opposition leaders cannot commit not to follow the same discriminatory strategy that the current ruler is using. This is the nature of the politics of fear: since current supporters of the leader are afraid of an ethnic switch in the leadership, they refrain from disciplining the leader except in the most excessive cases.

The amount the ruler is able to divert is endogenous to the model because both the gains from replacing her (avoiding her current stealing) and the costs (the lottery between a replacement from the current ruling group or the opposition) are determined in equilibrium by the strategies played by rulers and potential substitutes. This endogenization uncovers a mechanism of “amplification of kleptocracy”: current supporters will accept more stealing from their leader the bigger is the difference between their payoff under the current ruler and the payoff under a potential ruler from the opposition group. The more they could lose in the future (i.e. the bigger is the gap), the more their own ruler can steal from them today. This fear is greater the more the potential substitute from the opposition is able to steal in equilibrium. As a consequence, any characteristic of the economy that allows \textit{one} type of ruler to expropriate, will allow \textit{any} type of ruler to misbehave.

Several conclusions with respect to the policies and the amount of expropriation observed in equilibrium can thus be derived:

First, this logic of exclusion provides a rationale for the inefficient use of public funds as resources for patronage and redistribution to the ruler’s ethnic group. To the extent

\textsuperscript{14}African states are bureaucratically weak, which forces them to obtain resources via indirect taxation, or through manipulations of the price mechanism. See Van de Walle (2001) for a discussion on state capabilities.

\textsuperscript{15}Caselli and Coleman (2002) use the same idea in justifying the breakdown of societies in ethnic conflicts.
that each group’s income cannot be targeted specifically (if taxation is indirect, a group can escape a high rate of taxation by changing its economic activity) the ruler is forced to raise taxation rates across the board if she wants to increase the amount of funds at her reach. To prevent this increase from alienating her supporters, she needs to hand back targeted patronage to her group so that they keep supporting her. Since patronage is only received by a fraction of the population, but allows an increase of taxation to all groups in society, patronage will be overprovided in equilibrium. This outcome is thus consistent both with the coexistence of heavy rates of taxation and inefficient subsidization to supporter groups observed in Africa and with the prevalence of “pork-barrel” politics in ethnically divided societies.

Second, to the extent that different ethnic groups devote their efforts to different activities and there exist economic rigidities in their choice of activity (the presence of specific sources of wealth such as the qualities of the soil they till provides such rigidities) the ruler can use discriminatory taxation against the excluded group. Increased taxation on the activities of the excluded group will not be met by an immediate abandonment of such activities precisely because that group has a comparative advantage in them. In other words, the more a group’s source of wealth is sector specific, the less they can arbitrage differences of taxation across sectors. Hence, an economy based in long term cash crop exports such as cocoa or coffee, which involve long-term specific planted capital, will exhibit more venal leaders. The same holds for an economy richer in natural resources. This helps explain the natural resource curse of countries such as Nigeria, or the venality of Kenya’s leaders.

Both points above show that the excluded group can be discriminated against by the state both in taxation and in expenditure. However, the logic of the model makes clear that it would be misleading to conclude that there is redistribution from the excluded group to the dominant one. This pattern of discrimination is the result of the ruler’s maximization of resource expropriation from both groups. In other words, a significant share of the money she sends abroad for own consumption comes from her supporter’s pockets. This explains the ability of African leaders to amass such amounts of personal wealth.

Finally, the equilibrium amount of stealing is decreasing in the quality of institutions. In particular, institutions that make the capture of government by a particular group more difficult, reduce the personal dependence of political stability or limit the capacity

---

16 That there is stratification in economic activities across ethnic groups has been long noticed. "Cementing the ethnic division of labor is the preeminent role of ascriptive ties in economic relations in the developing world" in Horowitz (1985).
to target taxation and patronage will reduce kleptocratic excesses.

The implications of the model are thus entirely consistent with an array of cross-country empirical literature that concludes that ethnic fractionalization is associated with poverty and lost opportunities for growth\textsuperscript{17}. It results from these analyses that the main cause to link ethnic divisions to bad economic outcomes seems to be different dimensions of bad policy. A panoply of papers\textsuperscript{18} highlight the association of high levels of ethnic fractionalization with low provision of public goods such as infrastructure and publicly provided private goods such as schooling or health services. The picture is not only one of bad policies, but one of bad governance: ethnic fractionalization correlates positively with the presence of corruption, "pork barrel politics" and autocracy\textsuperscript{19}. A final stylized fact is that the relationship between ethnic fragmentation and negative economic outcomes is diluted at high levels of institutional development\textsuperscript{20}. In other words, ethnic divisions are more costly whenever institutions are weak\textsuperscript{21}.

The remainder of the paper is organized as follows. Next section relates this paper to different strands of theoretical literature. Section III presents the model and analyzes it, stressing the relevant comparative statics. It also contains a discussion on the interpretation of the model and its results. Section IV contains two extensions of the logic of the model. One explains another general pattern of public expenditure in Africa: the preponderance of wages over investment. A corollary on the relevance for ethnic violence is examined as well. Finally, the last section concludes.

## 2 Related literature

This paper is a contribution to the literature on political economy of less developed economies, in particular in the context of weak institutions. Ellman and Wantchekon (2000), La Ferrara and Bates (2001), Robinson and Verdier (2002), Robinson, Torvik and Verdier (2002) and Robinson and Torvik (2002) present models of electoral competition

\textsuperscript{17}This literature originates with Mauro (1995) and Easterly and Levine (1997), both papers using the same data to approximate ethnic divisions. The correlation of ethnic fractionalization with low levels of growth survives the use of alternative measures of ethnic divisions, as Alesina et al. (2003) show. This correlation is robust to the use of alternative summary indexes of diversity, such as polarization as in García-Montalvo and Reynal Querol (2002).

\textsuperscript{18}See Easterly and Levine (1997), Alesina et al. (1999),(2003) among others.

\textsuperscript{19}For the negative relationship of ethnic fractionalization with different indicators of civil liberties, electoral rights and democratic politics see Barro (1999) and Aghion, Alesina and Trebbi (2004). Mauro (1995), Alesina et. al (2003) among others show the positive relationship between fractionalization and different measures of corruption, bureaucratic inefficiency and absence of rule of law.


\textsuperscript{21}For a self-contained discussion of the relevant empirical findings as well as a survey of the literature see Alesina and La Ferrara (2003).
enriched to capture diverse characteristics of the political game in weakly institutionalized polities. In particular, the presence of ascriptive groups in society, the absence of commitment technology or the capacity to resort to violence are introduced to explain inefficient policy choices in these countries and the presence of clientelism. Even though my model is not explicitly electoral, it contributes to this literature by explaining why internal competition within the ruling group cannot dissipate kleptocratic rents even in a repeated game framework.

Other previous formal analyses of weak institutions include Acemoglu and Robinson (2004), Grossman (1991), Grossman and Noh (1994), Bueno de Mesquita et al. (2003) among many others. The general approach is to model the leader as maximizing the amount of resources she can extract from the polity subject to the constraint of remaining in power. I depart from previous work by capturing the idea that weak institutions include the opportunity to marginalize and expropriate part of the population. Moreover, I consider the consequences of non-institutionalized systems of succession and the particularities of ethnic divisions.

Acemoglu, Robinson and Verdier (2004) explicitly address the puzzle of the existence of kleptocratic rulers by noting that weak institutions imply that the leader can exacerbate the collective action problem of society. If the collaboration of different groups is needed to replace the ruler, she can use the existence of ruler-friendly institutions to buy off a pivotal actor and hence remain in power. The authors dub this strategy “divide-and-rule” following a tradition in the Political Science literature. While there is little doubt that collective action problems are important in explaining why autocratic leaders remain in power, my approach emphasizes a different mechanism that does not exclude the previous one, and explains why some groups actually explicitly give support to such kleptocratic leaders. In addition, my model generates inefficiencies along the equilibrium path, a property that Acemoglu et. al (2004) does not feature.

There is a small but growing formal literature on ethnic conflict. Caselli and Coleman (2003) provide a rationale for why conflicts tend to be along ethnic lines: characteristics such as language and skin color allow winners to exclude losers ex-post, hence changing the incentives to escalate conflict ex-ante. The idea of the role of excludability is thus present in their model, but my paper formalizes the idea that the elites are the ones contributing to and benefiting from ethnic divisions, and then explains why the rest of the population follows. To understand this process, it is crucial to separate the citizenry

---

22 The work of Bueno de Mesquita et al.(2003) already takes into account that members of the selectorate can be excluded from the winning coalition. Acemoglu et al. (2004) consider that the ability to tax different groups at different rates is a characteristic of weakly institutionalized polities.

from the leadership, a notion that formal models of group conflict do not tend to pursue\textsuperscript{24}.

3 The model

3.1 The Environment

Consider an infinitely repeated economy populated by a continuum of citizens of mass 1. Citizens belong to one of two ethnic groups, A and B. The size of group A is \( \pi^A \). In addition there are two economic activities, denoted a and b. A group is defined by two distinct sets of characteristics. First, there are some ascriptive characteristics like language or skin color (maybe geographical distribution) that are identifiable and, for simplicity, impossible to change\textsuperscript{25}. Second, each groups possesses a comparative advantage in a different portfolio of economic activities. A group A citizen obtains \( \omega^a \) per period in activity a. Should she decide to take activity b she would earn \( \omega^a - \theta^A \) per period. Symmetrically, a B citizen obtains \( \omega^b \) in activity b and \( \omega^b - \theta^B \) in activity a, per period. \( \theta^i \) captures the extent to which a group’s wealth is specific to a particular activity. For example, if a group obtains its wealth from coffee, it can switch its efforts to growing rice in their fields. Unfortunately, coffee trees are a long term specific investment, and hence putting those lands to another use entails a loss of pre-tax income. In general, a group that is specialized in cash-crops, especially tree crops, has no way to transfer its planted capital to another activity. The same would be true for a group that obtains revenue from natural resources that lie below its territory. On the other hand, \( \theta^i \) may simply capture the degree to which specialized knowledge is useless in another sector. Finally, a small value of \( \theta^i \) also captures the possibility that ethnic groups are not differentiated by economic activities. For simplicity, assume that switching is allowed each period. Let \( z^i_t = 1 \) if group i does not take the activity in which it enjoys comparative advantage in period t. Otherwise \( z^i_t = 0 \).

There is a state that performs two functions: it taxes economic activities and uses the proceeds to provide benefits to groups.

These benefits might be public goods that are so dependent on taste that only one of the groups enjoys them. The favoured interpretation is that they constitute pure patronage such as the allocation of public resources to the region of a group in the form

\textsuperscript{24}Some papers have appeared that make this distinction: Glaeser (2002) proposes a model in which hate is provided as a quasi-good by self-interested politicians to rational citizens who demand it. De Figueiredo and Weingast (1999) explain widespread violence and the participation of the masses as the reaction to uncertainty about the intentions of other groups and leaders.

\textsuperscript{25}In the spirit of Caselli and Coleman (2003)
of infrastructure or the granting of lucrative bureaucratic posts (or posts in the army, police, etc.) to members of the favored group. The state is able to discriminate across recipients for public expenditure thanks to the ascriptive characteristics of groups.

On the other hand, taxes are activity specific because in particularly poor developing countries as in Africa, the absence of a competent bureaucracy forces the governments to raise their revenue from indirect taxation. For instance, the use of Marketing Boards for agricultural products and other manipulations of the pricing system\textsuperscript{26} are pervasive. In the context of the model, I allow taxation to differ across group-activities, but the ability to imperfectly switch activities will put a ceiling on how differently one can tax different sources of wealth. Thus note that the fundamental difference between expenditure and taxation is that patronage can be perfectly targeted to specific groups.

At any point in time, one ethnic group has control of the government. Even though a group has the state nominally captured, real power is exercised by a narrow elite inside the group, and I will call it the Leader. Denote by $L^i$ the leader if she is from group $i$. In the remainder of the paper, I call the group to which the leader belongs "supporter" group, and the other is denoted "excluded" group for reasons that will become apparent. Each group has an unlimited supply of identical leaders from which to choose.

Denote $\tau^{ik}$ the tax level that a leader of group $i$ levies on activity $k$. Similarly, let $\eta^{ij}$ be the amount that leader of group $i$ spends on patronage for group $j$. Obviously $i,j \in \{A,B\}$ and $k \in \{a,b\}$. The amount $\eta^{ij}$ provides utility $R(\eta^{ij})$ to group $j$ with $R' > 0$, $R'(0) > 1$, $R'' < 0$ and $R(0) = 0$. Group $-j$ receives no utility from $\eta^{ij}$.

This economy has two fundamental states, $S_t \in \{A,B\}$, denoting whether power is captured by group $A$ or group $B$ in period $t$.

The instantaneous utility of a citizen of group $A$ in state $S$ (the expression for $B$ is just symmetric) is thus:

$$C(S, z^A) = (1 - z^A)(\omega^A - \tau^S a) + z^A(\omega^A - \theta^A - \tau^S b) + R(\eta^{SA})$$

where time subscripts have been omitted for notational simplicity.

Both groups have identical preferences represented by $E \sum_{t=0}^{\infty} \delta^t C^j_t$, where $C^j_t$ is the consumption of group $j$ at time $t$, and $\delta$ is the discount factor.

Even though the leader belongs to group $S_t$ she has self-serving interests. In particular, she wants to maximize the funds that she can divert for her own uses. A leader of group $A$ obtains instantaneous utility (the expression for $B$ is just symmetric) as long as she is

\textsuperscript{26}Bates (1981) provides a detailed account of these practices. In addition, Bates (1989) shows that these manipulations are inefficient to the point of contributing to famines.
in power:

\[ U^A = \pi^A (\tau^A - \eta^{AA}) + (1 - \pi^A) (\tau^{Ab} - \eta^{AB}) \]

and discounts future payoffs by \( \delta \). The expression assumes that there is no switching in equilibrium. When a leader is not in power, she obtains 0 utility per period.

The weakness of institutions and the importance of ascriptive links is captured in the model by the following assumptions. First, assume that whenever the incumbent leader retains the support of her kin group, she maintains her position with probability \( \bar{\gamma}^A \). With probability \( 1 - \bar{\gamma}^A \) group \( B \) is able to dislodge the leader from power and install a \( B \) leader even against a united \( A \) group. \( \bar{\gamma}^A \) might be well above \( \pi^A \) to capture the notion that weakness of institutions and the strength of ethnic links allow for a huge degree of incumbency advantage. This is the sense in which a group might be able to capture power. The unique credible source of support and thus the unique credible promise of future patronage is given by the ruler’s ethnic linkage with her own group\(^{27}\).

Second, if the supporters of an incumbent leader decide to subvert the authority of their leader and want to oust her from power, they succeed automatically. Hence the relevant constraint on the rapacious interests of the leader is the need to keep the support of her group. This is the sense in which the position of the leader is weak: she needs the active support of a sizable share of the population to maintain power.

Third, when a leader is ousted from power by her own supporters the state does not perform its functions for that period. Moreover, the group that is not in power will find it easier to use this opportunity to grab power and seat a leader from its ranks. This captures the reality of Personal Rule regimes in which successions are always uncertain matters, resolved in non-institutionalized ways. Thus, I assume that the status of the group in power will change with probability \( 1 - \bar{\gamma}^S \). \( \bar{\gamma}^S \) captures the degree to which the grip on power of group \( S \) is solid in independently of the personality of the ruler. In other words, \( \bar{\gamma}^S - \gamma^S > 0 \) captures the importance of “Personal Rule” since it measures the increased stability that retaining the incumbent buys to her supporters. On the other hand it could capture the entrenchment of the leader\(^{28}\) or the relative competence of the next replacement from the group.

The timing of each stage game, given state \( S_t \), is the following:

1. Leader \( L^S \) announces the policy vector \( P_t = \{ \tau_t^Sa, \tau_t^Sb, \eta_t^{SA}, \eta_t^{SB} \} \)

\(^{27}\)The group has to feature two characteristics: firstly, it has to be very costly to change one’s identity ex-post. Secondly, the ruler has to be able to commit to use the same basis of support in the future. The experience in Africa and elsewhere suggests that ethnic allegiances possess both features.

\(^{28}\)Removing an entrenched leader requires more significant subversion thus weakening the group vis-à-vis the challenging ethnic group.
2. The citizens of group $S_t$ decide to “subvert,” $s_t = 1$ or not, $s_t = 0$

3. If $s_t = 0$, the citizens decide $z_t^A$ and $z_t^B$ and afterwards the policy vector is implemented, payoffs are realized. Next period starts with $S_{t+1} = S_t$ with probability $\gamma^S$ and the state switches with probability $1 - \gamma^S$.

4. If $s_t = 1$, the leader is ousted immediately and the “revolt” vector $P_r = \{0, 0, 0, 0\}$ is implemented. With probability $1 - \gamma^S$, group $S$ loses power and the next period starts with $S_{t+1} = -S_t$. Otherwise, the next period starts with a new leader from group $S$.

There are a number of features of the model that are worth stressing. First, note that collective action within a group is not an issue in this model. The focus of the argument is on the forces that allow rents to be appropriated by a weak leader, and not competed away by different elites inside the same group. Adding heterogeneity and a collective action problem would only help the current leader to steal even more, because she would find it easier to disrupt coordination. Second, and in the same spirit, I do not allow the leader access to any repression instrument: if she loses the support of her group, she is replaced at no explicit cost.

Finally, note that no difference is made between democracy and dictatorship in the model. The evidence from Africa shows that democracies have not behaved differently than dictatorships at the time of supporting kleptocracies and corruption\(^\text{29}\). In my analysis, the reason is that both types of regimes have been able to play ethnic divisions and patronage networks in exactly the same ways. The analysis will reveal that institutional reform needs to go further than getting people to vote. It has to include effective constraints on the capacity of the leaders to treat ethnic groups differently, and it should include mechanisms directed to smooth intra-group competition.

### 3.2 Definition of Equilibrium

The equilibrium concept to be used is (pure strategy) Markov Perfect Equilibrium. In this type of equilibria, strategies can only be contingent on the payoff-relevant state of the world and the prior actions taken within the same period.

As has been described above, the state space of this economy includes only two elements, $\Theta = \{A, B\}$, denoting whether power is captured by group $A$ or group $B$ at the beginning of period $t$. Denote the state at each period by $S_t$, where obviously $S_t \in \Theta$.

\( \forall t = 0,1,2, \ldots \). Assume that each group has a set of potential leaders from which replacements will be drawn randomly. Call these two sets of leaders \( \Delta^A \) and \( \Delta^B \). At any point in time, the leader in power is denoted by \( L^A \) or \( L^B \) depending on the group she was drawn from. Denote by \( \nabla L^A \) the potential leaders that belong to \( \Delta^A \) but are not in power currently. \( \nabla L^B \) is defined symmetrically. The strategy of the current leader \( L^A \) is denoted by \( P^A \) and it is a four-tuple \( \{ \tau^{Aa}, \tau^{Ab}, \eta^{AA}, \eta^{AB} \} \in \mathbb{R}^4 \) when \( S_t = A \). When either \( S_t = B \) or \( S_t = A \) but a leader belongs to \( \nabla L^A \), her set of strategies is empty. The symmetric definition holds for the strategies of leaders \( L^B \).

The strategy of group \( A \) is denoted \( \sigma^A(S/P^S) \) and depends on both the state of political capture and the policy vector proposed by the leader. It determines two actions, \( \{ s^A, z^A \} \) that have been defined above as the decision to subvert and the decision to switch economic activities. If \( S_t = A, s^A \in \{0,1\} \), that is, if the leader is from group \( A \), this group can decide to give her support or to subvert her authority. On the other hand, if \( S_t = B \), \( s^A = \emptyset \). \( z^A \in \{0,1\} \) independently of the state. The symmetric definition holds for the strategy space of citizens of group \( B \).

State transitions work as follows: whenever \( s_t^S = 0 \), \( S_{t+1} = S_t \) with probability \( \gamma^S \) and the state switches with probability \( 1 - \gamma^S \). If \( s_t^S = 1 \), that is, if there is subversion, \( S_{t+1} = S_t \) with probability \( \bar{\gamma}^S \). Denote this transition function \( T(\sigma^S, S) \).

A (pure strategy) Markov Perfect Equilibrium for this game is a combination of strategies denoted by \( \{ \tilde{P}^A, \tilde{P}^B, \tilde{\sigma}^A, \tilde{\sigma}^B \} \) such that all four strategies are best responses to the other three for all possible states. In particular, consider the following set of Bellman equations:

\[
V^A(S) = \max_{\sigma^A} \{ C^A(S, \tilde{\sigma}^S, \sigma^A(S/P^S), \tilde{\sigma}^B) + \delta \sum_{S' \in \Theta} V^A(S') T(\sigma^S, S) \} \quad (1)
\]

\[
V^B(S) = \max_{\sigma^B} \{ C^B(S, \tilde{\sigma}^S, \sigma^B(S/P^S), \tilde{\sigma}^A) + \delta \sum_{S' \in \Theta} V^B(S') T(\sigma^S, S) \} \quad (2)
\]

\[
W^A_{L^A}(A) = \max_{P^A} \{ U^A(P^A, \tilde{\sigma}^A, \tilde{\sigma}^B) + \delta \sum_{S' \in \Theta} W^A_{\Delta}(S') T(\tilde{\sigma}^A(A/P^A), A) \} \quad (3)
\]

\[
W^B_{L^B}(B) = \max_{P^B} \{ U^B(P^B, \tilde{\sigma}^B, \tilde{\sigma}^A) + \delta \sum_{S' \in \Theta} W^B_{\Delta}(S') T(\tilde{\sigma}^B(B/P^B), B) \} \quad (4)
\]

where \( C^j \) denotes the consumption of citizen \( j \) as a function of the state \( S \) and the strategies of the leader in power and both sets of citizens. \( V^j(S) \) denotes the value function for citizen \( j \) in state \( S \). \( W^i_L(S) \) denotes the value function for leader from group \( i \) in state \( S \), when she is the current leader \( L^S \). To complete the definition, note that \( W^A_{\Delta}(B), W^A_{L^A}(A), W^B_{\Delta}(A) \) and \( W^B_{L^B}(B) \) are completely independent of any decision.
that the particular leader could take. They only depend on the probability that, in equilibrium, a particular leader will be in power in the future. As a consequence, these are not interesting strategic objects in this game. A Markov Perfect Equilibrium is thus a combination of strategies \{\hat{P}^A, \hat{P}^B, \hat{\sigma}^A, \hat{\sigma}^B\} such that \hat{\sigma}^A solves (1), \hat{\sigma}^B solves (2), \hat{P}^A solves (3) and \hat{P}^B solves (4).

### 3.3 Analysis

Assume without loss of generality that \( S_t = A \). The equilibrium is characterized by backwards induction within each stage game. Examine first the decision to switch the sector of production. Take first \( B \) producers. Note that the decision to switch does not affect continuation utilities, hence only the static difference in payoffs is relevant. After observing the policy vector \( P_t \), they will switch sector only if the loss in wealth is smaller than the difference in taxation. Formally,

\[
z_t^B = 1 \iff \omega^b - \tau^{Ab} < \omega^b - \theta^B - \tau^{Aa}
\]

Since it is in the interest of the ruler not to allow this switch, which is wasteful, this ability to switch provides an upper bound on the differential taxation that the ruler can levy on group \( B \). The effective constraint on the ruler will thus be

\[
\tau^{Ab} \leq \theta^B + \tau^{Aa}
\]

(5)

The equivalent restriction for group \( A \) is then

\[
\tau^{Aa} \leq \theta^A + \tau^{Ab}
\]

(6)

Obviously, both restrictions cannot be binding at the same time.

Let us examine now the decision to subvert by \( A \) supporters. Note that the leader is the first player to act in the stage game. As a consequence, since strategies can only be conditional on the state of the economy, a leader \( L^A \) always proposes the same policy vector \( P^A \). Upon observing \( P^A \), if there is no subversion \( (s_t = 0) \), \( A \) supporters obtain:

\[
\omega^a - \tau^{Aa} + R(\eta^{AA}) + \delta \gamma^A V^A(A) + \delta(1 - \gamma^A) V^A(B)
\]

Alternatively, if they subvert, \( s_t = 1 \), they expect:

\[
\omega^a + \delta \gamma^A V^A(A) + \delta(1 - \gamma^A) V^A(B)
\]
Hence the non-subversion condition reduces to:

\[ \tau^{Aa} - R(\eta^{AA}) \leq \delta(\gamma^A - \gamma^A)(V^A(A) - V^A(B)) \]  

(7)

Note that the ruler will always satisfy this constraint by subgame perfection. Not satisfying it gives her no benefit because in the period she is thrown out she already receives 0 utility, plus she will obtain 0 forevermore, while being in power implies receiving positive rents each period. Hence in any MPE there will never be any ousting of a ruler. Therefore the only possibility of a change of state is the excluded group wrestling power away from the supporter group, which happens each period with probability \(1 - \gamma^S\). In equilibrium the continuation values for a citizen A can thus be expressed as:

\[
V^A(A) = \omega^a - \tau^{Aa} + R(\eta^{AA}) + \delta(1 - \gamma^A)V^A(B)
\]

\[
V^A(B) = \omega^a - \tau^{Ba} + R(\eta^{BA}) + \delta\gamma^B V^A(B) + \delta(1 - \gamma^B)V^A(A)
\]

Solving these equations for \(V^A(A) - V^A(B)\) and substituting the resulting expression in (7), the no-subversion constraint can be written in terms of the equilibrium value of policy:

\[
\tau^{AA} - R(Z^{AA}) \leq \frac{\delta(\gamma^A - \gamma^A)}{1 + \delta(1 - \gamma^A - \gamma^B)}[\tilde{\tau}^{Ba} - R(\tilde{\eta}^{BA}) - \tilde{\tau}^{Aa} + R(\tilde{\eta}^{AA})]
\]

Where the superscript \(\sim\) denotes equilibrium values. For notational simplicity I will denote \(\Phi^A = \frac{\delta(\gamma^A - \gamma^A)}{1 + \delta(1 - \gamma^A - \gamma^B)}[\tilde{\tau}^{Ba} - R(\tilde{\eta}^{BA}) - \tilde{\tau}^{Aa} + R(\tilde{\eta}^{AA})]\). This term summarizes the way in which future equilibrium play affects present decisions. With these ingredients, now I am able to posit the problem of ruler \(L^A\):

\[
\max_{\{\tau^{Aa}, \tau^{Ab}, \eta^{AA}, \eta^{AB}\}} \pi^A(\tau_i^{Aa} - \eta_i^{AA}) + (1 - \pi^A)(\tau_i^{Ab} - \eta_i^{AB}) + \delta\gamma^A W_{LA}^A(A)
\]

(8)

subj.to

\[
\tau^{Ab} \leq \bar{B} + \tau^{Aa} \quad \text{[\(\lambda\)]}
\]

\[
\tau^{Aa} \leq \bar{A} + \tau^{Ab} \quad \text{[\(\nu\)]}
\]

\[
\tau^{Aa} - R(\eta^{AA}) \leq \Phi^A \quad \text{[\(\mu\)]}
\]

\[
0 \leq \eta^{AB} \quad \text{[\(\rho\)]}
\]

The ruler thus maximizes her returns per period, conditional on avoiding any wasteful
switching and subversion. The first order conditions of this program yield:

\[ \pi^A + \lambda - \nu - \mu = 0 \]  
\[ 1 - \pi^A - \lambda + \nu = 0 \]  
\[ -\pi^A + \mu R'(\eta^{AA}) = 0 \]  
\[ -(1 - \pi^A) + \rho = 0 \]

(9)  
(10)  
(11)  
(12)

The first order conditions are simple and easy to interpret. From (12) it is obvious that \( \eta^{AB} = 0 \). The reason is that providing patronage good to the excluded group is costly and yields no benefit, since what is critical is the support from the leader’s group. From (9) and (10) and the fact that \( \lambda \) and \( \nu \) cannot both be strictly positive at the same time we learn that \( \nu = 0 \), \( \lambda = 1 - \pi^A \) and \( \mu = 1 \). \( \nu = 0 \) implies that the second restriction is not saturated. The analysis thus reveals that the ruler endogenously chooses to discriminate against the "excluded" group. Quite intuitively the leader will tax the excluded group as much as she can, that is, to the point in which the first constraint is binding.

Every dollar that the ruler is able to tax her own supporters is worth more than one dollar for her, because it allows her to increase taxation on the excluded group. Note from (11) that \( \mu \) multiplies the return from the last unit of patronage given to group \( A \). The cost of this last unit is only \( \pi^A \), but its return is increased taxation from the whole population (because \( \mu = 1 \)). This disparity is the reason for inefficient patronage provision. The mechanism works as follows: the non-subversion constraint is binding and hence an increase in \( R(\eta^{AA}) \) allows the ruler to increase taxation on her supporters and, since the no-switching constraint is also binding, taxation on the excluded increases in parallel. Hence, patronage good for \( A \) is overprovided in equilibrium: the ruler considers the benefits from increasing taxation from the whole population, while a social planner would only consider the group that receives utility from it. This distortion is worse the narrower the basis of support of the ruler (the smaller \( \pi^A \)).

Formally, the stage program yields the following solution:

\[ \eta^{AB} = 0 \]  
\[ R'(\eta^{AA}) = \pi^A \]  
\[ \tau^{AA} = \Phi^A + R(\eta^{AA}) \]  
\[ \tau^{Ab} = \theta^B + \Phi^A + R(\eta^{AA}) \]

(13)  
(14)  
(15)  
(16)

The solution for patronage public goods (13) and (14) is thus independent of expectations of future play, but this is not the case for the amount of resources that the leader can
extract from both groups. In fact, the solution above presents a mapping between future equilibrium play and current taxation. Remember that, in equilibrium, another symmetric problem is solved by any $L^B$ leader in power. The solution to the program for $L^B$ is:

$$\eta^{BA} = 0$$
$$R'(\eta^{BB}) = 1 - \pi^A$$
$$\tau^{Bb} = \Phi^B + R(\eta^{BB})$$
$$\tau^{Ba} = \theta^A + \Phi^B + R(\eta^{BB})$$

(17)

(18)

Denote the mapping from expectations to current play $\Gamma(\Phi^A, \Phi^B) = (\tau^{Aa}, \tau^{Ab}, \tau^{Ba}, \tau^{Bb})$, given by (15), (16), (17) and (18). Moreover, the definition of $\Phi^A$ (and the symmetric definition of $\Phi^B$) provides a mapping from actual play to consistent expectations $\Psi(\tau^{Aa}, \tau^{Ab}, \tau^{Ba}, \tau^{Bb}) = (\Phi^CA, \Phi^CB)$. The equilibrium posits the requirement that these expectations be consistent with future play. In this context this reduces to finding a fixed point of the mapping that relates expectations into themselves: $\Psi(\Gamma(\Phi^A, \Phi^B)) = (\Phi^CA, \Phi^CB)$. Explicitly, this mapping is the following:

$$\Phi^CA = \frac{\delta(\tilde{\gamma}^A - \gamma^A)}{1 + \delta(1 - \tilde{\gamma}^A - \tilde{\gamma}^B)}[\theta^A + \Phi^B + R(\eta^{BB}) - \Phi^A - R(\eta^{AA}) + R(\eta^{AA})]$$

$$\Phi^CB = \frac{\delta(\tilde{\gamma}^B - \gamma^B)}{1 + \delta(1 - \tilde{\gamma}^A - \tilde{\gamma}^B)}[\theta^B + \Phi^A + R(\eta^{AA}) - \Phi^B - R(\eta^{BB}) + R(\eta^{BB})]$$

For simplicity denote $\Psi^i = \frac{\delta(\gamma^i - \tilde{\gamma}^i)}{1 + \delta(1 - \tilde{\gamma}^A - \tilde{\gamma}^B)}$. Solving this system for the fixed point $(\Phi^A, \Phi^B) = (\Phi^CA, \Phi^CB)$ yields:

$$\Phi^A = \frac{\Psi^A(1 + \Psi^B)(\theta^A + R(\eta^{BB})) + \Psi^A\Psi^B(\theta^B + R(\eta^{AA}))}{1 + \Psi^A + \Psi^B}$$

$$\Phi^B = \frac{\Psi^B(1 + \Psi^A)(\theta^B + R(\eta^{AA})) + \Psi^A\Psi^B(\theta^A + R(\eta^{BB}))}{1 + \Psi^A + \Psi^B}$$

Since there is a single fixed point, uniqueness of MPE is shown. This discussion establishes the following proposition.

**Proposition 1** The model presents a unique MPE. In equilibrium, in state $S = A$ (when $S = B$ the expressions are symmetric):
1. \( L^A \) proposes the following policy vector:

\[
\eta^{AA} \equiv \eta^*_A \text{ such that } R^r(\eta^*_A) = \pi^A \\
\eta^{AB} = 0 \\
\tau^{Aa} = \frac{\Psi^A(1 + \Psi^B)\theta^A + \Psi^A\Psi^B\theta^B}{1 + \Psi^A + \Psi^B} + (1 + \Psi^A)(1 + \Psi^B)R(\eta^*_A) + \Psi^A(1 + \Psi^B)R(\eta^*_B) \\
\tau^{Ab} = \frac{\Psi^A(1 + \Psi^B)\theta^A + (1 + \Psi^A)(1 + \Psi^B)\theta^B}{1 + \Psi^A + \Psi^B} + (1 + \Psi^A)(1 + \Psi^B)R(\eta^*_A) + \Psi^A(1 + \Psi^B)R(\eta^*_B)
\]

2. The citizens of group \( A \) accept this policy vector: \( s^A = 0 \)

3. No activity switch occurs: \( z^A = z^B = 0 \)

This MPE is related to the unique equilibrium of the finite horizon version of this game: denote by \( \Gamma_T \) a game with exactly the same stage game as the model presented above, but repeated a finite number of times \( T \). Since this game has a final stage, backwards induction can be used to find the unique subgame perfect equilibrium of the finite game. The following proposition establishes a link between this unique SPE and the unique MPE of the infinite horizon game.

**Proposition 2** Consider a game \( \Gamma_T \). Then:

1. \( \Gamma_T \) has a unique SPE

2. The limit as \( T \to \infty \) of the unique SPE of \( \Gamma_T \) yields the same payoffs as the MPE described in Proposition 1

The proof of this proposition can be found in the Appendix\textsuperscript{30}. This convergence exists because both equilibria are qualitatively identical: in both of them, the incumbent is never ousted, and she steals from her own supporters exactly the difference between staying under her rule and the lottery between replacements. It helps the intuition to examine what happens in the last period. The incumbent has to leave her supporters indifferent between ousting her and giving her support. Since there is no future, she

\textsuperscript{30}This result follows from uniqueness of SPE in the finite horizon game and uniqueness of MPE in the infinite horizon game (see Fudenberg and Tirole (1991)). Nonetheless, the proof in the appendix is constructive in the context of this game.
cannot threaten them with a bleak perspective and hence she has to give them at least $\omega^A$ utility. She does exactly this by taxing them $\tau^A(a) = R(\eta^A) = R(\eta^A)$ because this allows her to maximize extraction from the excluded group. Note that this creates a wedge between being in the supporter versus the excluded group in the last period that the leader can exploit in $T - 1$, $T - 2$. Hence, starting in $T - 1$, the leader can reduce her supporter’s utility below $\omega^A$. The further away from the end of the game, the less the constraint of giving $\omega^A$ to the supporters in the last period binds. Hence, the payoffs converge to the ones in the MPE from above\(^{31}\).

### 3.4 Discussion and Interpretation

The first subsection underscores a number of equilibrium characteristics and interprets them under the light of ethnic bias in Africa. The second one highlights the main lesson extracted from the game-theoretical modeling of the mechanism: any reason that allows one type of ruler to steal will spread throughout the economy to allow the venality of any type of ruler. The role of weak institutions is examined afterwards, shedding some light on the unwillingness of African leaders to develop their institutional framework. A final subsection explores the applicability of the specific mechanism developed in the model to different social and institutional frameworks than the ones pervasive in Africa.

#### 3.4.1 Policy Determination and Ethnic Bias

The unique MPE of the model provides an explanation for many features of the post-colonial political economy of Africa.

First, the model endogenously generates inefficient policies. Note that in the simple framework proposed here, the unique potential source of inefficiency is the excessive allocation of patronage to a particular group. Since the opportunity cost of public funds is 1, the fact that the marginal return to patronage for the supporter group is $\pi^A < 1$ shows that political needs cause inefficiencies. This feature of the equilibrium helps explain the patterns of inefficient taxation and inefficient transfers coexisting in the same group highlighted in the seminal work by Bates (1981) for agricultural policies in tropical Africa. The ruler needs to buy support from her own group while, at the same time, wants to

---

\(^{31}\)The model presents other SPE. For instance, citizens can use the MPE as a punishment device in trigger strategies equilibria in which they force the leader to present them with a particular level of utility, under the threat of being replaced. If the group that is supposed to replace does not do so, then the play jumps to the MPE forever. In any case, these equilibria need a lot of coordination behind the “veil of ignorance”, and a problem of divided societies is precisely that groups find it very difficult to trust, communicate and coordinate across ethnic lines. Avoiding dialog across groups is the basis of the “divide and rule” strategies played by the rulers.
extract a lot of resources from the economy. The best way of doing so, given the absence of lump-sum taxation, is by taxing both groups and then returning some patronage to the supporters even if this is highly wasteful. This is a general pattern of statism in Africa.

Second, the model predicts a very strong bias in the allocation of public funds. The excluded group receives no public benefits while the supporter group receives public resources beyond the optimal point. The use of public money in the form of bureaucratic posts, infrastructure or even access to schools as a form of patronage, as well as the ethnic bias in the allocation of these goods has been widely documented in Africa. Gikuyus and later Kalenjin in Kenya, northern groups in both Nigeria and Uganda or Tutsis in Burundi are just salient examples that reproduce across the continent. The bias in favor of the ruling group is conspicuous and is actually one of the basic sources of resentment between ethnic groups. Not only access to these positions is biased, but is accompanied by an absence of meritocratic pressure that makes them ripe for all kinds of corruption, official and unofficial.

Third, the bias is not only present in the allocation of patronage: taxation is also differential across groups. In particular, in addition to taxes levied on the supporter group, the model shows that the excluded group is expropriated from the non-transferable share of its wealth. Bates (1981) and Bates (1989) provide evidence of this pattern: in Ghana and Uganda, among other examples, the coalition that supported the leader extracted resources from the coffee and cocoa planters. These are crops that involve a lot of specific long term investment. On the contrary, in Kenya the Gikuyu controlled the coffee growing parts of the country, and hence the discrimination against these crops was much less evident.

The combination of higher taxation and absence of patronage makes the excluded group obviously worse off than the supporter group. As a consequence, whenever there is a change in the group controlling power, the patterns of taxation change and purges follow in order to make space for the new elites. For instance, the ascension to power by Moi in Kenya was followed by a substitution of Gikuyus by Kalenjin in all echelons of the state. In Ghana, cocoa has been heavily taxed by all governments, civil and military, except the one headed by Kofi Busia, a native from the Ashanti region which contains a large share of smallholders that grow cocoa. In Cameroon, the substitution of Ahidjo in 1982 unleashed another deep ethnic purge of the bureaucracy. Similar dynamics are found in Nigeria. Ironically, these purges tend to take place under the excuse of anti-corruption

---

33 See Collier and Garg (1999) for an account of how ethnic and kinship ties are rewarded in the form of salaries.
34 See Barkan and Chege (1989) for an account of the reallocation of posts and resources.
initiatives. These switches prove the use of public resources as patronage, as well as the conscious status of "ruling" groups versus "excluded" groups.

This pattern of discrimination both in raising revenue and in public expenditures supports the vision that a particular ethnic group has the government captured. The model suggests that the actual benefits of such capture are not spread throughout the group. The particular elite that holds power extracts so much resources that part of this money comes from the pockets of non-elite members of the group. In equilibrium it is very easy to see that (19) can be rewritten as $\tau^A = \Phi^A + R(\eta^A)$. Hence, since $\Phi^A > 0$, this is exactly the amount by which the ruler is able to reduce her followers’ utility. The next subsection studies the determinants of $\Phi^A$.

This result is also consistent with casual empiricism. In Kenya, a potential political cleavage, and a reason why Kenyatta used the ethnic card to maintain power was the situation of landless Gikuyus, most of them ex Mau Mau fighters. These downtrodden masses did not obtain anything from the regime, even though it was clear to all observers and political participants that Kenyatta was at the helm of a "Gikuyu" regime. Emphasizing the fact that the majority of Gikuyu were not actually receiving their share of the spoils was obviously threatening to the regime. That is why the leadership had to act hastily whenever any political entrepreneur tried to shed light on these facts. This included the assassination of a popular politician in 1975.

Wa Wamwere (2003) describes this absence of balance in the reception of spoils in a colorful way:

“The cream of government service goes to the ruling ethnic elites, the crumbs to the lesser ethnic elites, and dust to members of the so-called ruling ethnic community” and “Among the Gikuyu of Kenya, the approving masses are called grill lickers, njuna ndara”.

The fact that non-elites are not receiving much from the government is by no means unique to Kenya.

Fourth, the results of the model rationalize the existence of kleptocratic elites supported by masses of impoverished ethnic followers. Even though in absolute terms the masses are made worse off by the existence of rent-creating policies, in relative terms it is

---

35 These dynamics have been crystallized in the much repeated sentence "it is our turn to eat". See, for example, Wa Wamwere (2003).

36 See Throup and Hornsby (1998:19). They write that this leader was “(...) attempting to mobilize the kikuyu masses -the masakini (literally, the poor)- against the conspicuous wealth of the kikuyu elite, especially Kenyatta’s relatives and close allies.”

37 “Ordinarily it is the representatives and fiduciaries of ethnic groups, more than the general members, who gain privileges or suffer punishments under systems of personal rule,” from Jackson and Rosberg (1983)
much better to belong to the group in power than to the excluded group, and hence they are willing to defend the status quo vis à vis a leader from another group. The members of a narrow elite around the leader are thus the ones extracting the lion’s share of the rents that these inefficient policies create. Evidence of Kleptocratic tendencies abound in Africa, but Mobutu’s Zaire is probably the most cited example. Sani Abacha in Nigeria or Daniel arap Moi in Kenya have been able to amass personal fortunes counted in the billions of dollars\textsuperscript{38}. Even a relatively well-considered leader, such as Houphouet-Boigny in Cote d’Ivoire had his share of personal aggrandizement projects, such as a marble covered cathedral in his home town. Consistent with this concentration of wealth at the highest levels of leadership, Africa is the continent with highest capital flight\textsuperscript{39}.

3.4.2 Amplification of Kleptocracy

The theoretical reason that supports kleptocratic regimes in this model is summarized in expression (7). It makes clear that as long as the supporter group observes a difference between being in the supporter status and being excluded under the leadership of the opponent group, there is a surplus that the current leader can expropriate from her own supporters. In addition, the more a leader can extract from her supporters, the more she can extract from the excluded group, thanks to (5) being binding in equilibrium. As a consequence, there is an amplification effect of any characteristic of the economy that allows one ruler to steal. Assume that the institutional or economic technology of this society changes so that $L^B$ is now able to steal more from her group if she is ever in power. An $A$ citizen understands that, in equilibrium, this will mean that should he ever fall into an excluded status, his plight will be worse. This reduces $V^A(B)$ in equilibrium. But obviously, this losses the non-subversion constraint for $L^A$ and as a consequence, $L^A$ is able to increase $\tau^{Aa}$ to the point where her supporters are again indifferent between giving her support or subverting and taking a lottery that now is much less favorable, since both $V^A(A)$, and $V^A(B)$ are reduced. This amplification mechanism is the reason why in the expressions for equilibrium taxation in Proposition 1 the economic and institutional characteristics of both groups appear.

Expression (19) can be rewritten to identify the substantive forces that allow the leader to reduce the utility of her own group:

$$\tau^{Aa} = \frac{\Psi^A(1 + \Psi^B)\theta^A + \Psi^A\Psi^B\theta^B}{1 + \Psi^A + \Psi^B} + \frac{\Psi^A\Psi^BR(\eta^A_s) + \Psi^A(1 + \Psi^B)R(\eta^B_s)}{1 + \Psi^A + \Psi^B} + R(\eta^A_s)$$  \hspace{1cm} (21)$$

\textsuperscript{38}See Ayittey (1992), Wa Wamwere (2003), Mbaku (2000) or any account of corruption in Africa.

\textsuperscript{39}Collier and Gunning (1999)
The gap between $\tau^A_A$ and $R(\eta^A_A)$ is exactly $\Phi^A$, the amount the ruler reduces her supporters’ utility. It is easy to see that the forces that allow leaders to create a wedge between supporters and excluded are their ability to discriminate in taxation, captured by the first summand and their capacity to allocate patronage, captured by the second summand in (21). Hence, both tools that the ruler can manipulate in this reduced form model are used. If the ethnic structure does not coincide with an economic sectorial cleavage or there are no important specificities in the economy (this would be a case in which $\theta^A$ and $\theta^B$ are small) the ruler cannot discriminate in taxation but patronage still makes a difference in the utility of her supporters and hence they are still reluctant to increase the probability of an ethnic switch in the leadership.

In the model, the net amount of funds that the leader $L^A$ is able to extract equals $X^A = \Phi^A + R(\eta^A_A) + (1 - \pi^A)\theta^B - \pi^A\eta^A_A$. By the envelope theorem (and because $\mu = 1$ in (8)), all the interesting effects enter through $\Phi^A$:

$$\frac{\partial X^A}{\partial \theta^A} = \frac{\partial \Phi^A}{\partial \theta^A} = \frac{\Psi^A(1 + \Psi^B)}{1 + \Psi^A + \Psi^B} > 0$$

$$\frac{\partial X^A}{\partial \theta^B} = \frac{\partial \Phi^A}{\partial \theta^A} + 1 - \pi^A = \frac{\Psi^A\Psi^B}{1 + \Psi^A + \Psi^B} + 1 - \pi^A > 0$$

The amount of expropriation from both groups, independently of the allegiance of the leader, is increasing in the share of non-transferable resources in the economy. This result implies that starting from a situation with low $\theta^A$ and $\theta^B$, an increase in the share of non-transferable wealth anywhere in the economy increases equilibrium misbehavior by the ruler. In the case of a citizen of group $A$, and increase in $\theta^A$ implies increased expropriation by a potential $L^B$. As a consequence, she allows her leader to steal more from her. An increase in $\theta^B$ has two effects: the direct one comes from the tax markup that $L^A$ charges on group $B$. In addition, there is the amplification effect: an increase in $\theta^B$ means the $B$ citizens will be afraid of losing power if they ever regain it, and hence a $B$ leader will be able to steal more from them and, as a consequence, steal more from $A$ citizens that would be excluded in that case. Therefore an increase in $\theta^B$ allows $L^A$ to reduce her supporters’ utility further. These comparative statics provide a rationale for the well documented "natural resource curse": an increase in $\theta^i$ may capture the discovery of some mineral resource in the land of a particular group.

Comparative statics with respect to the ethnic demographic balance are ambiguous. On the one hand, all the direct effects predict a reduction in stealing: increasing $\pi^A$ reduces the benefits from distorting the patronage good for two reasons. First, rents are reduced at each level of provision because it becomes more expensive to provide it.
Moreover, the optimal level of distortion is reduced because the returns are reduced (less people in the excluded group to pay for it).

In addition, increasing $\pi^A$ reduces the fraction of population excluded, and hence reduces the extra revenue that comes from the extraction of their non-transferable resources.

However, there is a third, indirect effect, that makes the overall effect ambiguous: increasing $\pi^A$ means that, should the group ever lose power, a potential $L^B$ would be able to steal more: she would additionally distort the allocation of $\eta^B$ because her basis of support would now be smaller. Using the same logic of amplification, this allows an $L^A$ leader extra room for stealing. Explicitly, the partial derivative has the following expression:

$$\frac{\partial X^A}{\partial \pi^A} = \frac{\Psi^A \Psi^B R'(\eta^A)}{1 + \Psi^A + \Psi^B R''(\eta^A)} - \eta^A - \theta^B - \frac{\Psi^A (1 + \Psi^B) R'(\eta^B)}{1 + \Psi^A + \Psi^B R''(\eta^B)}$$

The first two summands represent the rents lost from the ability to distort $\eta^A$ and the third is the direct loss that is a consequence of the smaller size of the excluded group. The last summand represents the indirect effect, and it is positive. For general functional forms of $R(\cdot)$ this expression cannot be signed, but note that if $R(\cdot)$ is a power function, $\frac{R'(\eta)}{R''(\eta)}$ is increasing in $\eta$. Hence, if the third indirect effect ever dominates, it will do so at high levels of $\pi^A$. That is, when the $A$ group includes a wide majority of the population, the prospect of falling under an $L^B$ is most terrifying because she will have a very narrow basis of support, and hence she will use extreme distortions of patronage to steal.

### 3.4.3 The Effect of Institutions on Rent-Seeking

If the probability of an ethnic turnover captures the degree of institutional strength, it is informative to analyze the comparative statics of stealing with respect to these set of parameters. In particular, it is easy to show that:

$$\frac{\partial X^A}{\partial \gamma^A} = \frac{\partial \Phi^A}{\partial \gamma^A} = \frac{\partial \Phi^A \partial \Psi^A}{\partial \Psi^A \partial \gamma^A} + \frac{\partial \Phi^A \partial \Psi^B}{\partial \Psi^B \partial \gamma^A} > 0$$

$$\frac{\partial X^A}{\partial \gamma^B} = \frac{\partial \Phi^A}{\partial \gamma^B} = \frac{\partial \Phi^A \partial \Psi^A}{\partial \Psi^A \partial \gamma^B} + \frac{\partial \Phi^A \partial \Psi^B}{\partial \Psi^B \partial \gamma^B} > 0$$

$$\frac{\partial X^A}{\partial \gamma^A} = \frac{\partial \Phi^A}{\partial \gamma^A} = \frac{\partial \Phi^A \partial \Psi^A}{\partial \Psi^A \partial \gamma^A} + \frac{\partial \Phi^A \partial \Psi^B}{\partial \Psi^B \partial \gamma^A} < 0$$

$$\frac{\partial X^A}{\partial \gamma^B} = \frac{\partial \Phi^A}{\partial \gamma^B} = \frac{\partial \Phi^A \partial \Psi^A}{\partial \Psi^A \partial \gamma^B} + \frac{\partial \Phi^A \partial \Psi^B}{\partial \Psi^B \partial \gamma^B} < 0$$

\(\text{Recall that it is assumed that } R'' < 0\)
None of these comparative statics are ambiguous. From these comparative statics it is clear that the level of rent extraction is increasing in both \((\gamma^A - \bar{\gamma}^A)\) and \((\gamma^B - \bar{\gamma}^B)\). The first one is obvious: the leader can extract more resources from her followers the more their probability of keeping power depends on maintaining this particular leader. The fact that \(L^A\) can expropriate more the bigger \(\gamma^B - \bar{\gamma}^B\) is follows from the logic of amplification of kleptocracy discussed in the previous subsection.

Hence stealing increases with institutional uncertainty and personality-dependent control of power. In particular, the leader would like to reduce the grip on power of her followers if she is ousted, while at the same time strengthen her ethnic group position vis-à-vis the excluded group as long as she is in power. While these parameters depend on characteristics of the polity beyond the control of the ruler, such as the demographic ethnic balance, it certainly also depends on institutional factors. In particular, it has been shown that uncertainty in the succession process and incumbency advantage bought via clientelist networks is a characteristic of systems of personal rule.

These comparative statics explain the weakening of the correlation between ethnic fractionalization and bad economic outcomes at high levels of institutional development. Even with a divided society a leader cannot extract much from the citizenry if the hold in power of a particular group and the stability of a particular regime does not depend on the personal links of the ruler on top. This reduction on the dependency on personalities is precisely a sign of institutional strength and hence it is not surprising that the margin to misbehave in strongly institutionalized polities is very much reduced.

This model is not the right framework to think about questions of institutional development but the logic of the mechanism shows that the leader has no incentive to strengthen the institutional framework if this means increasing her accountability. This sheds light on several facts.

First, this is consistent with the behavior of the leadership in most African countries: from the moment of independence, even the first prophetic leaders such as Nkrumah, clamped down on opposition, banned political parties, used the police and the military in a partisan way, did not respect judicial independence or any kind of separation of powers and imposed censorship on the press. These are not actions of rulers interested in institutional consolidation. These actions repress both excluded groups and potential replacement leaders from the group of the current leader, thus increasing \(\bar{\gamma}^i - \gamma^i\).

Second, it is very important for the leaders not to allow the presence of a strong and obvious second-in-command. This would permit her followers to replace her and quickly coordinate on giving support to this alternative focal point and thus reduce the risk

\[^{41}\text{The expressions for each partial derivative are listed in Appendix 2 for completion.}\]
of being taken over. In the context of the model, this would increase \( \gamma^i \). Jackson and Rosberg (1982) write "As long as a ruler retains command in African states, an overriding consideration in succession rivalries is that they be concealed from him". The reason given in their book is that presenting candidacy as successor may signal the politician as an ambitious man keen on substituting the ruler, and hence a potential target for oppression. The reason in the model proposed in this paper is more general. Even if the successor could guarantee not to plot against the ruler, his mere designation reduces institutional uncertainty and, as a consequence, reduces the margin of the leader to steal. A typical way of getting rid of close collaborators that have become too powerful or focal, and therefore a potential threat is the claim of their involvement in a coup plot against the current leader. Note that this is also a pervasive feature of regimes of personal rule.

3.4.4 Out of Africa

This mechanism is especially relevant in explaining the extent of kleptocracy, clientelism and inefficiency in Africa because of the continent’s societal characteristics and the political constraints which its leaders face, but by no means it is exclusive to politics in Africa. To extrapolate the argument to other circumstances, the key ingredients of the mechanism have to be described in detail.

First, existence of groups is necessary. Two characteristics of group definition are essential: it has to be very difficult to cross the group division ex-post, and it has to be possible to specifically target the utility of the groups with policy. The saliency of ethnicity in Africa creates a natural rigid dividing line. However, a strong ideological divide would suffice. For instance, a citizen can classify herself as a moral conservative, and she knows that tomorrow she will most probably still be a moral conservative. Some dimensions of policy, such as religious and moral tolerance affect her utility differently from the utility of a liberal person. For example, Divorce Legalization probably increases the utility of a socially liberal but reduces that of a moral conservative.

Second, the group-identity of the leader has to be able to create a difference in the payoff of groups. The absence of institutional constraints enlarges the set of policies that the African ruler can use to widen the utility gap between supporters and opposition. In particular, an African leader can use violence, discrimination and patronage to increase the utility of her supporters and reduce that of the excluded. As institutional

---

42There are numerous accounts of the obsession of long term leaders in refusing to designate anybody as a successor, being Malawi’s Banda probably the most notorious.

43Sékou Touré’s Guinea was famous for the ridiculous amount of coup plots allegedly discovered against the leader. See Jackson and Rosberg (1982) or Cartwright (1983) for an account of these facts.
strength increases, the leader faces a reduction in her set of feasible discriminatory policies. Nonetheless, the identity of the leader still makes a difference in issues such as social policy and foreign policy which can target ideological groups differently even in strongly institutionalized democracies. Given the group-identity of the ruler, she is able to commit to these policies in the future and, as a consequence, enlist the support of her group.

Third, support from the group must make a difference in maintaining the leader in power. Arguably, in an African Personal Rule regime, the ability to mobilize ethnic allegiances is essential for the survival of the leader. Moreover, since succession mechanisms are not established, uncertainty follows subversion. However, to some extent, every voter in a strongly institutionalized democracy faces a similar dilemma: the only way for the voter to punish a perceived deviation\textsuperscript{44} from a leader from her group is not to vote for her. Unfortunately, by withholding the vote, the voter is marginally increasing the probability that a candidate from the opposite group is elected, which entails a loss in utility.

Hence, the difference between a well-functioning democracy and the kleptocracies that are present in Africa is just one of degree: strong institutions limit the extent to which a group can capture power and restrict the set of tools that the leader can use to widen the gap between supporters and the opposition. As a consequence, stealing or shirking in office is restrained. As institutions place less constraints on the leader and societal divisions grow wider, accountability of the rulers is weakened. In the case of Africa, both circumstances occur in their extreme form and, as a consequence, outright kleptocracy is sustainable.

4 Extensions

4.1 On Public Investment

The evidence from Africa shows that governments overspend in wages and undertake very little of infrastructure construction\textsuperscript{45}. The intuition from the model clarifies why this is the case.

Note that in the model wages are explicitly considered. The nature of patronage $\eta$ is such that it is a flow concept that has to be pledged at every stage and has no consequences for the future. Public wages fall into this category: bureaucrats can easily

\textsuperscript{44} Strong institutions affect the possibility of outright stealing in western countries. But there are many instances of ideological shirking or other ways in which the ruler extracts a personal benefit at a cost for her supporters.

\textsuperscript{45} See Collier and Gunning (1999) for an account of these patterns. See, as well, Mbaku (2000) and Bates (1981) for some examples in Cameroon, Nigeria and Ghana of excess expenditure in wages.
be replaced or moved to different areas. As a consequence, by expanding public sector employment beyond optimality among her supporters and placing them in their area, the leader creates a network of clients personally invested on the continuation of the status quo. This is a reason besides outright corruption and misappropriation why public sector wages swallow a much bigger share of public expenditures in Africa than in other less developed economies.

A simple extension of the model that includes the possibility to invest in durable infrastructures clarifies which kind of projects suffer from underinvestment when the leader can use divisions and weak institutions to capture the support of her group.

Assume that the policy vector \( P_t^S = \{\tau_{Aa}^t, \tau_{Ab}^t, \eta_{AA}^t, \eta_{AB}^t, I^t\} \) is now a 5-tuple that includes \( I \), public investment. This investment contributes to a stock of public capital \( K \). This stock evolves according to the following dynamics: \( K_t = \varsigma K_{t-1} + I_{t-1} \). Hence, investment today increases the stock of public capital tomorrow and this capital depreciates at a rate \( 1 - \varsigma \). The stock of infrastructure provides a benefit \( F(K) \) to the supporter group and \( \beta F(K) \) to the excluded group, for \( \beta \leq 1 \). \( \beta \) thus captures the degree of excludability of public infrastructure. A pure public good would have \( \beta = 1 \). Assume that \( F' > 0 \) and \( F'' < 0 \). If there is replacement of a leader, the revolt vector is extended to \( P_r = \{0, 0, 0, 0, 0\} \) and therefore the government does not invest for that period. Assume, finally, that when there is a revolt, the enjoyment of the public infrastructure is reduced to \( \psi F(K) \), for \( \psi \leq 1 \). \( \psi \) captures the instantaneous cost of upheaval.

To simplify the analysis, and in particular the dynamics that a new state variable could introduce, assume that leaders have no financial constraints and their instantaneous utility is linear: \( U_t^A = \pi_A (\tau_{Aa}^t - \eta_{AA}^t) + (1 - \pi_A) \tau_{Ab}^t - (K^A_{t+1} - \varsigma K^A_t) \). This implies that whatever is the optimal level of capital in steady state, it will be reached as soon as a leader has a chance to invest. As a consequence, the transitional period that replacing the leader induces lasts only one stage, precisely because of the absence of investment in the presence of a revolt. Hence I can set the problem in terms of the desired level of capital for next period \( \tilde{K}^A \), and investment will just be determined as a residual: \( I_t = \tilde{K}^A - \varsigma K^A_t \). In addition, assume that \( \hat{\gamma}^A = \hat{\gamma}^B = 1 \), for expositional clarity (in this case, support from the group is enough to keep power with certainty).

The model can be solved in exactly the same way as the previous case\(^{46}\). In particular, the Markov Perfect Equilibrium has the same characteristics: the leader in power satisfies the constraint that makes her supporters indifferent between replacing her and supporting

\(^{46}\)The inclusion of a new state variable may create multiplicity of MPEs. For expositional ease I only analyze the equilibrium most similar to the previous case in which strategies are not conditional on the stock of capital.
her rule. Hence, in equilibrium there is no replacement of the leadership. To examine explicitly the new support constraint, note that when supporters do not revolt, $s_t = 0$, they receive:

$$\omega^A - \tau_{t}^{Aa} + R(\eta_t^{AA}) + F(K_t^{A}) + \delta(\omega^A - \tau_{t+1}^{Aa} + R(\eta_{t+1}^{AA}) + F(K_{t+1}^{A})) + \delta^2 V^A(A).$$

On the other hand, if they revolt, $s_t = 1$, they obtain:

$$\omega^A + \psi F(K_t^{A}) + \delta(\gamma^A[\omega^A - \tau_{t+1}^{Aa} + F(\varsigma K_t^{A})] + (1 - \gamma^A)[\omega^A - \tau_{t+1}^{Ba} + \beta F(\varsigma K_t^{A})]) + \delta^2(\gamma^A V^A(A) + (1 - \gamma^A)V^A(B)).$$

Note that in writing these expected values I am already making use of the fact that the transitional period will only last one stage before setting into a new equilibrium. Hence, the support constraint, using stationarity, can be written as:

$$\tau_t^{Aa} - R(\eta_t^{AA}) + (\psi - 1)F(K_t^{A}) + \delta(1 - \gamma^A)[\tau_{t+1}^{Aa} - R(\eta_{t+1}^{AA}) - \tau_{t+1}^{Ba}] - \delta F(K_{t+1}^{A}) + \delta(\gamma^A + (1 - \gamma^A)\beta)F(\varsigma K_t^{A})$$

$$\leq \frac{\delta^2}{1 - \delta}(1 - \gamma^A)[\tau^{Ba} - \beta F(\tilde{K}^{B}) - \tau^{Aa} + R(\tilde{Z}^{AA}) + F(\tilde{K}^{A})]$$

The ruling leader faces a problem identical to (8) with an additional choice variable, $K_{t+1}^{A}$ and the support constraint replaced by (22). The first order conditions from this program imply that the chosen level of $K_{t+1}^{A}$ will be determined implicitly by the following expression:

$$F'(K_{t+1}^{A})(2 - \psi) - \delta[\gamma^A + (1 - \gamma^A)\beta]F'(\varsigma K_{t+1}^{A}) = \frac{1}{\delta} - \varsigma$$

The interpretation of this expression is a little involved. It can be separated in two parts: $F'(K^{A})(1 - \psi)$ is the marginal effect on $\tau^{Aa}$ caused by the contemporaneous effects of upheaval: if there is no revolt, supporters enjoy the whole return from infrastructure, while subversion reduces it to a fraction $\psi$. $F'(K^{A}) - \delta[\gamma^A + (1 - \gamma^A)\beta]F'(\varsigma K^{A})$ is the effect on $\tau^{Aa}$ caused by the effect of upheaval on next period’s returns: if there is no subversion, citizens will receive the full return per period, $F(K^{A})$ while ousting the leader has two effects. First, it will reduce the stock of capital tomorrow to $\varsigma K^{A}$, because of the absence of investment today. Second, the citizens enter on the lottery for the replacement,
which means that their expected enjoyment is scaled down by \([\gamma^A + (1 - \gamma^A)\beta]\). Note that the equivalent expression for a social planner that would take the welfare of both groups into account would be:

\[
[\pi^A + \beta(1 - \pi^A)]F'(K_{t+1}^A) = \frac{1}{\delta - \varsigma}
\]

By comparing the two expressions we can obtain a sense of which parameter values will make them diverge more. First, the right hand side of both conditions is equal: marginal costs are the same for both the social planner and the ruler, and obviously, costs increase with depreciation.

Second, the left hand side of (23) implies that the leader will invest more in capital the faster the good depreciates and the less useful it is during upheaval. That is, her investment is decreasing\(^{47}\) in \(\varsigma\) and \(\psi\). These two forces are ignored by the social planner\(^{48}\).

There are political reasons behind these effects: reduction in \(\pi\) and/or \(\psi\) increases the costs of subverting the authority of the leader. In particular, one of the costs of upheaval is the absence of investment for one period, because the state stops working. Obviously, this is not a large cost if depreciation is very slow and, as a consequence, investment is very small each period. On the contrary, when \(\varsigma\) approaches 0, \(K\) behaves very similarly to \(\eta\), and it has been shown above that in this case there is overspending in patronage.

The reason for investment to decrease with \(\psi\) is exactly the same.

Third, excludability (\(\beta\)) features in both expressions with contrary effect. When supporters know that they can be easily excluded from enjoying public capital if there is a change in the leadership, they are more eager to defend the current regime. Hence an increase in \(\beta\) reduces investment by the leader. For the social planner, investment is increasing in \(\beta\) because more people are able to enjoy the public infrastructure and social welfare increases while costs remain the same.

Assuming that \(F(K) = K^\alpha\), for \(\alpha < 1\), allows an explicit look at the gap between the capital level of the leader and the socially optimal level. Denote by \(\tilde{K}^A\) the level chosen by the ruler and by \(K^*\) the level chosen by the social planner. The ratio of both expressions satisfies:

\[
\left(\frac{\tilde{K}^A}{K^*}\right)^{1-\alpha} = \frac{2 - \psi - \delta[\gamma^A + (1 - \gamma^A)\beta]\varsigma^\alpha}{\pi^A + \beta(1 - \pi^A)}
\]

Note that this ratio is decreasing in \(\beta\), \(\psi\) and \(\pi\). The intuitions have been explained above.

Note that when \(\beta = \psi = \varsigma = 1\), the relative level of capital provided by the leader is very

\(^{47}\)As long as \(\varsigma F'(\varsigma K)\) is increasing in \(\varsigma\).

\(^{48}\)Obviously, if the good depreciates faster any investor would invest less in it, but this is the effect present in the right hand side of both expressions.
low, because $\delta$ is close to 1.

The absence of investment in infrastructure can thus be interpreted under the light of the model. The political survival of the leader hinges on creating a network of supporters personally dependent on her presence in power. Building a road gives no advantage to the leader, because the day a coup ousts her from power the road will still be there for everybody to enjoy. The same is true for a hospital or a school. Hence there is underprovision of capital when it is very permanent, it is not excludable and/or does not lose usefulness during periods of upheaval. At the other extreme of the parameter space, there is overprovision of capital. This is why the leader favors expenditures in wages and reduces infrastructure construction\footnote{This argument for the existence of inefficient clientelist networks based on "flow" goods such as expanded public employment versus "stock" goods such as infrastructure is examined explicitly in Robinson and Verdier (2002). They develop other comparative statics, but the framework presented here explicitly links the appearance of these networks with the weak institutions and ethnic divisions present in Africa.}

### 4.2 On Patterns of Ethnic Violence

The logic of the model can be extended to show that another way the leader can enlist the support of her ethnic group is by making sure that ethnic supporters disproportionately fear the prospect of being under the rule of another group. The ruler can contribute to this fear by acting heavy-handedly against oppressed groups, and making sure that her rule is seen by everyone as ethnically based rule. Ayittey (1999) describes how ordinary Krahn people feared the demise of Samuel K. Doe, a fellow Krahn, in Liberia. Even though they did not receive any of the spoils from government, the fact that the regime was clearly almost exclusively Krahn, and that it was engaged in acts of pillaging, rape and atrocities against the other groups made clear to them that retribution would be against all Krahn the day the regime was defeated. Obviously, this made even more Krahn people collaborate in defending the regime. Hence, ethnic violence can be used to enlist otherwise reluctant members of the group in the defense the regime.

To see how the prospect of violence is equivalent to patronage links to the ruler the model can be extended in a very simple way. Assume that by oppressing the excluded group, the leader can contribute to the level of enmity that the excluded group holds against the supporter group. Call $E_t^{AB}$ the level of enmity that group $B$ has against $A$. Assume that this variable evolves in the following way $E_t^{AB} = \phi E_{t-1}^{AB} + o_t^{AB}$, where $o_t^{AB}$ denotes the amount of costly oppression that $L^A$ inflicts over $B$ citizens. Assume further that living under the leadership of a group that stocks enmity against you causes disutility because of revenge, which is captured by $\Pi(E^{AB})$, with $\Pi' > 0$, $\Pi'' < 0$, and
Π(0) = 0.

Now, if supporters don’t subvert, \( s_t = 0 \), they will receive\(^50\):

\[
\omega^A - \tau_t^{Aa} + R(\eta_t^{AA}) + \delta V^A(A)
\]

If they replace the leader, \( s_t = 1 \), they will receive:

\[
\omega^A + \sum_{t=1}^{\infty} \delta^t \{ \bar{\omega}^A(\omega^A - \tilde{\tau}^{Aa}) + R(\eta^A_*) \} + (1 - \bar{\gamma}^A)(\omega^A - \tilde{\tau}^{Ba} - \Pi(E^A_s)) \}
\]

Hence, the support constraint can be written as:

\[
\tau_t^{Aa} - R(\eta_t^{AA}) \leq (1 - \bar{\gamma}^A) \sum_{s=t+1}^{\infty} \delta^s(-\tilde{\tau}^{Aa} + R(\eta^A_*) + \tilde{\tau}^{Ba} + \Pi(E^A_s)) \tag{24}
\]

Hence, the program of the leader is to maximize \( \sum_{t=1}^{\infty} \delta^t \{ \pi^A(\tau_t^{Aa} - \eta_t^{AA}) + (1 - \pi^A)\tau_t^{Ab} - (E^A_{t+1} - \phi E^A_t) \} \) under the usual no-switching constraints and (24). The optimal level of enmity that \( L^A \) seeks (as in the previous subsection she will jump right away to this level because of linear utility) is determined by:

\[
\sum_{s=t+1}^{\infty} (\delta \phi)^s \Pi(\phi^s E^A_s) = \frac{1 - \phi \delta}{1 - \bar{\gamma}^A}
\]

This expression shows that the leader will cause a higher level of enmity the slower enmity disappears (the higher \( \phi \)), the smaller the time discounting (the higher \( \delta \)) and the bigger the chance that the supporters will lose control of power should they replace the leader (the lower \( \bar{\gamma}^A \)). The intuition is perfectly in line with the rest of the argument developed in the paper: \( V^A(B) \) is smaller the slower the rate of forgiveness. Moreover, \( V^A(B) \) is more probable when \( 1 - \bar{\gamma}^A \) is small. Both effects make her supporters warier of entering into the replacement lottery and this allows the ruler to extract more resources.

To see that enmity behaves similarly to the capacity to discriminate in patronage, rename \( \Lambda^{AB} = (1 - \bar{\gamma}^A) \sum_{s=1}^{\infty} \delta^s \Pi(\phi^s E^A_s) \). With this change of notation the model can be solved for the unique MPE which has the following expression for taxes extracted from the supporter group:

\[
\tau^{Aa} = \frac{(1 + \zeta^B)\zeta^A \theta^A + \zeta^B \zeta^A \theta^B}{1 + \zeta^A + \zeta^B} + (1 + \zeta^B)\zeta^A A^{BA} + (1 + \zeta^B)\Lambda^{AB} + (1 + \zeta^A)(1 + \zeta^B)R(\eta^A_*) + \zeta^A(1 + \zeta^B)R(\eta^B_*)}{1 + \zeta^A + \zeta^B}
\]

\(^50\) Again, for simplicity the case shown assumes \( \bar{\gamma}^A = \bar{\gamma}^B = 1 \)
where $\zeta^A = (1 - \gamma^A) \frac{\delta}{1 - \delta}$ and $\zeta^B = (1 - \gamma^B) \frac{\delta}{1 - \delta}$. Note that in this expression, the capacity to induce enmity enhances the capacity to steal in a similar way as the ability to discriminate in patronage. Kleptocracy can thus be supported by indiscriminate use of violence in the name of the group, as Samuel K. Doe’s Liberia show.

This logic can be used to understand the scale of atrocities and ethnic cleansing in Rwanda in 1994. Prunier (1995) describes how the "hutu" regime of Habyarimana and the inner circle captained by his wife was besieged both by the Tutsi guerrillas of the RPF and the southern moderate hutu elites that were complaining at the level of corruption and kleptocracy concentrated in northern hands. By making the majority of the hutu population participate in the atrocities, the regime almost succeeded in doing two things. First they completely erased the northern-southern divide inside the hutu elites by either eliminating those hutus deemed too moderate or making them participate in the genocide. Second, the scale of atrocities against the tutsi minority was so horrific, that no hutu could accept the prospect of living under a tutsi leader for fear of equally horrible retribution. The massive scale of hutu refugee tides to Zaire is a testament to this strategy. This pattern of government sponsored ethnic violence, albeit in a somewhat smaller scale has been present in Uganda, Burundi51 and other countries in the region.

Post-colonial rulers have not been the first ones in using these strategies in Africa. Multiple academic accounts, among them Cooper (2002) and Horowitz (1985) describe the process by which ethnic separation became a basic strategy of domination by the colonial powers. Some groups where protected and allowed to thrive but at the same time they were demonised among the rest of the population. In reality, political reasons were paramount in deciding whether nilotes were good for the army in Uganda and baganda were good for civil service. This process contributed dramatically to the creation of ethnic self-consciousness and resentment but generated a basis of support for colonial presence. These societies were, as a consequence, ripe for the exploitation of such divisions by their post-colonial leaders.

The logic of exclusion and replacement thus provides as a corollary a framework to understand some of the patterns of high and low level ethnic violence that plague deeply divided societies, especially when their governments define themselves in ethnic terms.

51 See Lemarchand (1996) for an account of the seeds of violence and patterns of ethnic domination in Burundi.
Post-colonial African leaders have provided very bad outcomes to their populations. This has not so much been the result of economic incompetence but of flagrant abuse of power to impose distortionary and rent-creating policies on their economies. Moreover, these rents have been dissipated at the very top of the leadership, and some countries have been governed by outright kleptocracies. Nonetheless, accounts coincide in considering these regimes weak, so it is not credible that they have survived in power thanks to the use of force to oppress the whole population.

This paper advances a different explanation for such blatant absence of political accountability: weak institutions have allowed leaders to exploit ethnic divisions. In particular, the use of patronage networks to treat the ethnic group of the leader better than the opposition makes ethnic supporters keen in maintaining this relative superiority. If weak institutions have allowed the leader to establish a system of personal rule, replacing the ruler will be accompanied by upheaval and consequentially the current opposition could take control of the state. This combination makes supporters reluctant to replace the leader and hence rents are not dissipated: the ruler is able to extract resources from all groups in society. From the analysis I derive a number of corollaries that relate the amount of funds diverted by the leader to the structure of the economy and the quality of institutions.

Obviously, there are a number of factors contributing to the absence of accountability that have been left out of the stylized model. Here I name some of them, for further research. First, a particularly interesting factor is the collective action problem. Ousting a leader that governs in a personal rule regime probably entails activities that imply personal risk. Since the benefit of replacing the leader is collective, it is obvious that there is scope for free-riding. I did not include this element in the model because I wanted to show that even in the absence of collective action problems these leaders are able to establish kleptocracies in deeply divided societies, but there is no doubt that there is a lot to be learnt from an explicit analysis. In particular, oppression of the press has been a typical feature of these regimes. If there is a collective action problem, not allowing the press to report freely can contribute to make coordination matters even worse.

Second, it would also be interesting to model explicitly the role of violence in keeping a sizable share of the population in the excluded status, as well as oppressing the emergence of any potential leader among the supporter group. These two patterns of violence seem, for theoretical as well as empirical reasons quite different and an explicit model may show, for example, why these regimes have kept the military weak while at the same time have
used ethnic militias as a close, albeit difficult to control, substitute.

Third, the model assumes a society divided in the most trivial way: two groups. Adding more groups, as well as allowing sizes to vary may offer some insights on the type of ethnic dictatorship that will emerge. It would be interesting to see if the existence of "pivotal" groups makes kleptocracy a less viable alternative, even though the empirical evidence does not seem to support this hypothesis. Is there an ethnic configuration of society that is differentially conducive to abusive rule?

Fourth and common with the majority of models that treat ethnic groups as given, the insight offered begs the constructivist critique: would this conclusion hold if we allowed groups to be generated endogenously by the action of ethnic elites? I would argue that not only the conclusion would be robust, but that we would actually observe these leaders trying to create or activate cleavages in society, following the lines of the colonizers. However, this process is not modeled explicitly and from such an engagement other lessons may be learnt.

The main lesson of the growing field of analysis of policy determinants in weakly institutionalized regimes is that their inefficiencies emerge because weak states impose a very particular set of contraints and needs on their rulers. The same is true for the political outcomes of such countries. Attempts at helping these economies have to take into account where the incentives of their leaders are (mis)placed. In particular, given the absence of accountability highlighted here, it is not surprising that enormous amounts of foreign aid have passed unnoticed through African economies to end up stashed in Swiss bank accounts. This paper is an attempt to offer a theoretical foundation for those who insist that serious institutional reform has to be attempted to solve the plight of African citizens.
References


Appendix 1: Proof to proposition 2

First, note that a finite horizon version of the game will have a unique subgame perfect equilibrium (SPE): there are no simultaneous moves in the tree and instances of indifference are the result of the usual open set argument. Hence the generically unique SPE can be found by backwards induction.

Denote by $V_{T}^{AA}$ the present discounted utility of the unique SPE for citizen $A$ if the game starts with a leader $L^A$ and the stage game is repeated $T$ times. $V_{T}^{AB}$ and $V_{T}^{BA}$ are defined equivalently.

Assume, without loss of generality that the game ends with a leader $L^A$. It is very easy to see that in the final stage of the game, in the unique SPE, the leader offers the following policy vector: $\{\tau_{T}^{AA}, \tau_{T}^{AB}, \eta_{T}^{AA}, \eta_{T}^{AB}\} = \{R(\eta_{*}^{A}), R(\eta_{*}^{A}) + \theta_{B}, \eta_{*}^{A}, 0\}$, where $\eta_{*}^{A}$ is such that $R'(\eta_{*}^{A}) = \pi^{A}$. In this fashion the leader is providing utility $\omega^{A}$ to her supporters while optimally stealing from group $B$. The reason why she cannot reduce her supporters’ utility is that there is no future and hence she would be replaced if her proposed policy induced a reduction in citizens’ $A$ utility. As a consequence, the leader is not replaced in the last period (her supporters are indifferent) and she earns some positive rents. Hence, the unique SPE of any finite horizon game will have the same properties as the MPE of Proposition 1: the leader never loses support and she steals to the point that her supporters are indifferent. Specifically, at each stage, the leader maximizes:

\[
\max_{\{\tau_{t}^{AA}, \tau_{t}^{AB}, \eta_{t}^{AA}, \eta_{t}^{AB}\}} \sum_{t=1}^{T} \delta^{t-1}\left[\pi^{A}(\tau_{t}^{AA} - \eta_{t}^{AA}) + (1 - \pi^{A})(\tau_{t}^{AB} - \eta_{t}^{AB})\right] \quad \text{[25]}
\]

subject to:

\[
\tau_{t}^{AB} \leq \theta_{B} + \tau_{t}^{AA} \quad \text{[25a]}
\]

\[
\tau_{t}^{AA} \leq \theta^{A} + \tau_{t}^{AB} \quad \text{[25b]}
\]

\[
V_{T}^{AA} \geq \omega^{A} + \delta \gamma^{A} V_{T-1}^{AA} + \delta(1 - \gamma^{A}) V_{T-1}^{BA} \quad \text{[27]}
\]

\[
0 \leq \eta_{t}^{AB} \quad \text{[28]}
\]

Constraints (25) and (26) are the usual no-switching constrains, while constraint (27) is the non-stationary finite horizon no-subversion constraint. In particular, the left hand side is the utility the supporters receive if they do not subvert (that is, they play the equilibrium strategy for a game of length $T$) and the right hand side is the expected utility of subverting: no state interference today and the lottery between leaders from tomorrow onwards which provides exactly the same utility as a game of length $T-1$.

The solution to this program yields that (25), (27) and (28) are binding at every $t$. 

41
State the definition of the present discounted utility in equilibrium:

\[
V_{T}^{AA} \equiv \omega^A - \tau_1^{Aa} + R(\eta_s^A) + \delta \tilde{\gamma}^AV_{T-1}^{AA} + \delta(1 - \tilde{\gamma}^A)V_{T-1}^{BA} \\
V_{T}^{AB} \equiv \omega^B - \tau_1^{Ab} + \delta \tilde{\gamma}^AV_{T-1}^{AB} + \delta(1 - \tilde{\gamma}^A)V_{T-1}^{BB}
\]

By substracting the second equation from the first and using (25) as binding, one obtains an equation in differences:

\[
V_{T}^{AA} - V_{T}^{AB} = \omega^A - \omega^B + \theta^B + R(\eta_s^A) + \delta \tilde{\gamma}^A(V_{T-1}^{AA} - V_{T-1}^{AB}) - \delta(1 - \tilde{\gamma}^A)(V_{T-1}^{BB} - V_{T-1}^{BA})
\]

Note that the same equation can be written in a symmetric fashion from the point of view of \( L^B \). Now, denote \( \Theta_T^A = V_{T}^{AA} - V_{T}^{AB} \) and \( \Theta_T^B = V_{T}^{BB} - V_{T}^{BA} \). The two version of the previous equation can be written as a system in differences:

\[
\begin{pmatrix}
\Theta_T^A \\
\Theta_T^B
\end{pmatrix} =
\begin{pmatrix}
\delta \tilde{\gamma}^A & -\delta(1 - \tilde{\gamma}^A) \\
-\delta(1 - \tilde{\gamma}^B) & \delta \tilde{\gamma}^B
\end{pmatrix}
\begin{pmatrix}
\Theta_{T-1}^A \\
\Theta_{T-1}^B
\end{pmatrix} +
\begin{pmatrix}
(\omega^A - \omega^B + \theta^B + R(\eta_s^A)) \\
(\omega^B - \omega^A + \theta^A + R(\eta_s^B))
\end{pmatrix}
\]

The two eigenvalues of the transition matrix of this system are \( \delta \) and \( \delta(\tilde{\gamma}^A + \tilde{\gamma}^B - 1) \) hence both are smaller than 1 in absolute value. As a consequence the system converges to a well defined \((\Theta^A_{\infty}, \Theta^B_{\infty})\).

Using the definition of \( \Theta_T^A \) and \( \Theta_T^B \), the system formed by (27) and the equivalent restriction from the problem of \( L^B \), taken as binding, can be written as:

\[
V_{T}^{AA} = \omega^A + \delta \tilde{\gamma}^A V_{T-1}^{AA} + \delta(1 - \tilde{\gamma}^A) V_{T-1}^{BB} - \delta(1 - \tilde{\gamma}^A) \Theta_{T-1}^B \\
V_{T}^{BB} = \omega^B + \delta \tilde{\gamma}^B V_{T-1}^{BB} + \delta(1 - \tilde{\gamma}^B) V_{T-1}^{AA} - \delta(1 - \tilde{\gamma}^A) \Theta_{T-1}^A
\]

Which can be rewritten as a dynamic system:

\[
\begin{pmatrix}
V_{T}^{AA} \\
V_{T}^{BB}
\end{pmatrix} =
\begin{pmatrix}
\delta \tilde{\gamma}^A & \delta(1 - \tilde{\gamma}^A) \\
\delta(1 - \tilde{\gamma}^B) & \delta \tilde{\gamma}^B
\end{pmatrix}
\begin{pmatrix}
V_{T-1}^{AA} \\
V_{T-1}^{BB}
\end{pmatrix} +
\begin{pmatrix}
\omega^A - \delta(1 - \tilde{\gamma}^A) \Theta_{T-1}^B \\
\omega^B - \delta(1 - \tilde{\gamma}^B) \Theta_{T-1}^A
\end{pmatrix} \quad (29)
\]

(29) is the general expression for the payoffs of a game of length \( T \), starting the iteration with \( V_1^{AA} = \omega^A \) and \( V_1^{BB} = \omega^B \). The eigenvalues of this transition matrix are \( \delta \) and \( \delta(\tilde{\gamma}^A + \tilde{\gamma}^B - 1) \) and hence they are both smaller than 1. Since the autonomous additive term converges, it is easy, but very tedious to show, that the payoffs converge to the invariant vector of this mapping with \( \Theta^B_{\infty} \) and \( \Theta^A_{\infty} \). The expression for the invariant
vector is thus:

\[
\begin{pmatrix}
V^A_A \\
V^B_B
\end{pmatrix}
= \begin{pmatrix}
1 - \delta \gamma^A & -\delta(1 - \gamma^A) \\
-\delta(1 - \gamma^B) & 1 - \delta \gamma^B
\end{pmatrix}^{-1}
\begin{pmatrix}
\omega^A - \delta(1 - \gamma^A)\Theta^B_\infty \\
\omega^B - \delta(1 - \gamma^B)\Theta^A_\infty
\end{pmatrix}
\]

After some lengthy algebra, it turns out that:

\[
\begin{pmatrix}
V^A_A \\
V^B_B
\end{pmatrix}
= \frac{1}{1 - \delta}
\begin{pmatrix}
\omega^A \\
\omega^B
\end{pmatrix}
- \frac{1}{(1 - \delta)(1 + \delta(1 - \tilde{\gamma}^A - \tilde{\gamma}^B))}
\left((1 - \delta\tilde{\gamma}^B)\Phi^A + \delta(1 - \tilde{\gamma}^A)(\Phi^B - \theta^A - R(\eta^B))\right)
\left((1 - \delta\tilde{\gamma}^A)\Phi^B + \delta(1 - \tilde{\gamma}^B)(\Phi^A - \theta^B - R(\eta^A))\right)
\]

Which are exactly the payoffs for the unique MPE of the infinite horizon game presented in Proposition 1.
7 Appendix 2: Comparative Statics

The expressions for the comparative statics of $\Phi^A$ with respect to the institutional parameters depend on the following partial derivatives:

$$\frac{\partial \Phi^A}{\partial \Psi^A} = \frac{(1 + \Psi^B)^2(\theta^A + R(\eta^{BB})) + (1 + \Psi^B)\Psi^B(\theta^B + R(\eta^{AA}))}{[1 + \Psi^A + \Psi^B]^2} > 0$$

$$\frac{\partial \Phi^A}{\partial \Psi^B} = \frac{\Psi^A(\theta^A + R(\eta^{BB})) + (1 + \Psi^A)\Psi^A(\theta^B + R(\eta^{AA}))}{[1 + \Psi^A + \Psi^B]^2} > 0$$

$$\frac{\partial \Psi^A}{\partial \bar{\gamma}^A} = \frac{\delta(1 - \delta \bar{\gamma}^B) + \delta^2(1 - \gamma^A)}{[1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)]^2} > 0$$

$$\frac{\partial \Psi^A}{\partial \gamma^A} = -\frac{\delta}{1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)} < 0$$

$$\frac{\partial \psi^A}{\partial \bar{\gamma}^A} = \frac{\delta^2(\bar{\gamma}^A - \gamma^A)}{[1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)]^2} > 0$$

$$\frac{\partial \psi^A}{\partial \gamma^A} = 0$$

$$\frac{\partial \psi^B}{\partial \bar{\gamma}^B} = \frac{\delta^2(\bar{\gamma}^B - \gamma^B)}{[1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)]^2} > 0$$

$$\frac{\partial \psi^B}{\partial \gamma^B} = 0$$

$$\frac{\partial \psi^B}{\partial \bar{\gamma}^A} = \frac{\delta(1 - \delta \bar{\gamma}^A) + \delta^2(1 - \gamma^B)}{[1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)]^2} > 0$$

$$\frac{\partial \psi^B}{\partial \gamma^A} = -\frac{\delta}{1 + \delta(1 - \bar{\gamma}^A - \bar{\gamma}^B)} < 0$$