A model of equilibrium institutions[*]

Bernardo Guimaraes[†]
Sao Paulo School of Economics (FGV)
& London School of Economics

Kevin D. Sheedy
London School of Economics

First draft: 19th October 2009
This draft: 30th September 2010

Abstract

In order to understand inefficient institutions, one needs to understand what might cause the breakdown of a political version of the Coase Theorem. This paper develops a model of power and distribution where institutions (the "rules of the game") are set to maximize payoffs of those individuals in power. They are constrained by the threat of rebellion, where any rebels would be similarly constrained by further threats. Equilibrium institutions are the fixed point of this constrained maximization problem. Private investment depends on credible limitations on expropriation, which can only be achieved if power is not as concentrated as those in power would like it to be, ex-post. Endogenously, this enables the group in power to act as government committed to protection of property rights, which would otherwise be time inconsistent. But the "political" Coase Theorem does not hold. Since sharing power implies sharing rents, capital taxation is inefficiently high.

JEL classifications: D7; H1; H2; O1; P4.

Keywords: government; political economy; public goods; capital taxation; time inconsistency.

*PRELIMINARY. We thank Bruno Decreuse, Erik Eyster, Carlos Eduardo Goncalves, Ethan Ilzetzki, Per Krusell, Dirk Niepelt, Andrea Prat, Ronny Razin and seminars participants at the Anglo-French-Italian Macroeconomic Workshop, U. Carlos III, Central European University, Columbia University, CREI conference "The political economy of economic development", ESSIM 2010, Institute for International Economic Studies, London School of Economics, Paris School of Economics, U. Sao Paulo, Sao Paulo School of Economics/FGV, U. Surrey, and U. St. Gallen for helpful comments.

†Corresponding author: Sao Paulo School of Economics (FGV), Email: bernardo.guimaraes@fgv.br.

And Samuel told all the words of the LORD unto the people that asked of him a king.

And he said, This will be the manner of the king that shall reign over you: He will take your sons, and appoint *them* for himself, for his chariots, and *to be* his horsemen; and *some* shall run before his chariots. ...

And he will take your daughters *to be* confectionaries, and *to be* cooks, and *to be* bakers. ...

He will take the tenth of your sheep: and ye shall be his servants.

And ye shall cry out in that day because of your king which ye shall have chosen you; and the LORD will not hear you in that day.

Nevertheless the people refused to obey the voice of Samuel; and they said, Nay; but we will have a king over us;

That we also may be like all the nations; and that our king may judge us, and go out before us, and fight our battles.

1 SAMUEL 8:10–20

1 Introduction

Institutions are defined by North (1990) as the "rules of the game", or the "humanly devised constraints that shape human interaction". Differences in institutions go a long way in explaining the huge disparities in income across the globe as they affect incentives for agents to invest, produce and exchange.[1] Even though institutions may serve the interests of elites, a fundamental question is why this gives rise to huge economic inefficiencies, that is, why is there no "political" equivalent of the Coase Theorem?

This paper builds a model to address that question. The model starts from a world of ex-ante identical individuals. The group in power (the *elite*) and the rules laying down the allocation of resources will both be endogenously determined in equilibrium, with the goal of maximizing the payoffs of those who will be in the elite. The elite and the rules comprise the equilibrium institutions.[2]

The only constraint on the choice of institutions is the threat of "rebellions", which destroy the prevailing institutions. There are no other "technological" constraints on which institutions

[1]For example, see North and Weingast (1989), Engerman and Sokoloff (1997), Hall and Jones (1999) and Acemoglu, Johnson and Robinson (2005).

[2]Recently, some models have been developed aiming at understanding institutions. Greif (2006) combines a rich historical analysis of trade and institutions in medieval times with economic modelling, part of which focuses on the form of government and political institutions that emerged in Genoa. Acemoglu and Robinson (2006, 2008) analyse conditions leading to democracy or dictatorship in an environment where an elite is trying to maintain its power, while citizens prefer a more egalitarian state. In Besley and Persson (2009a,b, 2010), society comprises two groups of agents that alternate in power, and make investments in two technologies that respectively allow the state to tax people and to enforce contracts. The exogenous parameters are the extent of political turnover and institutional (or demographic) features that determine how much one group cares about the other. They obtain predictions consistent with the data on state capacity and fiscal capacity, civil wars, and different forms of taxation.

are feasible, thus the issue will be the incentives for creating and destroying institutions. In the absence of rebellion, the rules prescribed by the chosen institutions are followed. This means that an elite can establish institutions laying down credible rules as long as it avoids rebellions. So although institutions will be chosen in the interests of an elite, this does not preclude the emergence of a "political" Coase Theorem. But crucially, once institutions have been destroyed, the new institutions that will arise cannot be constrained by the past rules or any other deals made earlier. This reflects the commitment problems analysed in Acemoglu (2003) and Acemoglu, Johnson and Robinson (2005).

Rebellions are costly, reflecting the fact that conflict is costly. Importantly, anyone can participate in a rebellion, including members of the current elite. Moreover, there is no limit to the number of rebellions: after any group takes power, there is always another opportunity for rebellion. The cost of conflict for rebels is proportional to the number of individuals in power who do not take part in the rebellion.

Since individuals are ex-ante identical and self-interested, the notion of equilibrium institutions is independent of the competence, benevolence, or factional affiliation of the individuals comprising the elite. Hence, once a rebellion has succeeded in destroying the current institutions, the rebels will have the same objectives and face the same constraints as those formerly in power. As in George Orwell's *Animal Farm*, there is no intrinsic difference between the "men" and the "pigs", but in equilibrium, some individuals will be "more equal" than others.

In the model, it is not conflict per se, but rather the threat of conflict that shapes institutions by constraining the actions of those in the elite. No conflict occurs in equilibrium owing to the absence of both randomness in the conflict technology and uncertainty about the actions of individuals.

The incentives to rebel depend on what members of a new elite would be able to extract once in power, which is also true of rebellions against institutions set up by the rebels, and so on. Given that all individuals are ex-ante identical, there is no fundamental reason why different institutions would be chosen following a rebellion, so the paper focuses on Markovian equilibria. The equilibrium institutions are thus the fixed point of the constrained maximization problem of the elite in power subject to the threat of rebellion, where the rebels would be similarly constrained by further threats of rebellion.

There is no restriction on the composition of a group launching a rebellion. Since the most disgruntled individuals have the most to gain from a rebellion, discouraging the "most profitable" rebellion a fortiori discourages all possible rebellions. It follows that the equilibrium institutions assign payoffs to producers according to a maximin rule, and hence, payoffs of those outside the elite (the "producers") are equalized. Furthermore, as the survival of institutions depends on those in the elite not defecting and joining a rebellion, payoffs of those in power must also be equalized. In other words, sharing power implies sharing rents.

In an endowment economy, there is a basic trade-off that characterizes the equilibrium institutions. The larger the elite (and hence the greater the number of people with a stake in defending the institutions), the greater the amount of taxes that can be levied, but the proceeds need to be divided among more people. The problem can be represented as a choice of the size of the elite and taxes to maximize the payoff of a member of the elite, subject to the constraint of avoiding a rebellion by producers. This "no-rebellion" constraint acts as a participation constraint for the producers.

Do the equilibrium institutions lead to the elite acting as a "government" in any meaningful sense of the term? That is, are institutions that are designed to maximize the payoffs of those in power ever congruent with the interests of the people? In some cases, the answer turns out to be yes. For example, suppose there is a technology that transforms the output of the producers into a public good that benefit everyone, and which has no impact on any other aspect of the environment. Such a public good will be optimally provided in equilibrium, as if it had been chosen by a benevolent government.

This outcome is consistent with a world characterized by a political Coase Theorem. The ability of the elite to set down rules is analogous to the possibility of contracting in the "regular" Coase Theorem. The notion of the elite constrained by the threat of rebellion means that there is a "price" attached to policy actions affecting those outside the elite, analogous to the existence of markets in the Coase Theorem.

A natural application of the model is to the taxation of investment proceeds. The model is extended so that once institutions are formed and after no rebellion has occurred, workers have access to an investment technology. However, the fruits of this investment are realized only after a lag, during which time there is another round of opportunities for rebellion in which both workers and members of the elite can participate. Since the cost of investing is sunk, it does not affect individuals' incentives to rebel, so by the maximin principle for payoffs in equilibrium explained earlier, the group in power would have incentives to expropriate fully individuals' investments.

Although institutions could in principle prescribe any level of capital taxation, whatever is chosen must be "rebellion-proof" after investment decisions have been made. The problem is that the capital arising from investment increases incentives for attacks on the current institutions that lead to new rules permitting full capital expropriation. Thus it is necessary to raise the cost of rebelling. This can only be done by expanding the size of the elite. The problem cannot be solved through taxes and transfers as such instruments can only redistribute disgruntlement with the institutions, decreasing rebellion incentives for some, but raising them for others.

In equilibrium, to ensure there will be no incentives to rebel against the existing institutions — leading inevitably to full expropriation — the elite has to be large so that if a rebellion were to take place after investment decisions had been made, the equilibrium size of the subsequent elite would be smaller than the current one. This is the only way to prevent members of the elite simply launching a costless rebellion from within that maintains intact the elite's composition (a rebellion is costless if the entire elite is willing to defect). Following a rebellion, there would then be insufficient places for all the current elite in the subsequent one. Thus, some members of the elite will oppose changing the institutions, and so conflict with them makes it costly to expropriate capital.

Adding the possibility of investment to the model thus gives rise to a larger elite. Sharing power among a wider group of individuals allows the elite to act as a government committed to a certain set of policies that would otherwise be time inconsistent. The model highlights the importance of sharing power as a way to guarantee stability of institutions and thus incentives for investment. This resonates with Montesquieu's doctrine of the separation of powers, which is now accepted and followed in all well-functioning systems of government. It is important to note that in the model, power is not shared among those individuals who are actually investing. The extra individuals in the elite in no sense represent or care about those who invest — but they do care about their own rents under the status quo. Thus, this group of self-interested individuals acts a government that commits to some protection of property rights.

Although it is possible to sustain protection against expropriation in equilibrium, capital taxation is set far from efficiently, so the political Coase Theorem breaks down. While in general it would be optimal from the point of view of society to have a larger group in power in order to guarantee that the fruits of investment would not be expropriated, the equilibrium elite chooses taxes on investment that are too high.[3]

There are two reasons for the inefficiently high level of captital taxation. First, the elite cannot extract all surplus from investors as the effort required to invest is private information. In equilibrium, the payoff of those who invest is larger than the payoff of the workers who do not invest, so the no-rebellion constraint for the investors is not directly binding.

The second (and more interesting) distortion follows from the distributional effects of protecting against expropriation. Since members of the elite can bring down the institutions more easily if they take part in a rebellion (because in that case they won't defend the current institutions), they must receive rents. As lower capital taxes require more power sharing, they also imply more rent sharing. This goes against the interests of each individual in the elite.

These reasons allow us to understand why a political Coase Theorem breaks down in this case. The first reason mentioned above is the unobservability of investors' surpluses. The second reason for the failure of the Coase Theorem is essentially the inseparability of power

[3]This is in accordance with the empirical literature that highlights the importance of institutions guaranteeing protection against government expropriation. For example, the results in Acemoglu and Johnson (2005) suggest that such institutions are more important than those that facilitate contracting among private agents.

and rents: it is not possible in equilibrium to add individuals to the elite and grant them a payoff lower than their peers. That leads to an endogenous limitation (which is binding) on the set of possible transfers.

While the model is quite abstract, it is congruent with a number of historical examples, some of which are discussed later in the paper: the disappearance of private corporations (the *societas publicanorum*) when power was concentrated under the Roman emperors; the need for a militarily strong leader (*podestà*) to guarantee stability in a society (medieval Genoa) where other strong groups could seize power; the tenacious resistance of the Stuart Kings of England to sharing power with Parliament.

Section 2 presents the model of power and distribution. The benchmark provision of public goods is briefly analysed in section 3, and the case of private investment in presented in section 4. Section 5 draws some conclusions.

1.1 Related literature

Since Downs (1957) emphasized the importance of studying governments composed of self-interested agents, a vast literature on political economy has developed (see, for example, Persson and Tabellini, 2000). Most of this literature focuses on democracies, so institutions are not themselves explained in terms of the decisions of self-interested agents. But in much of the developing world and during most of human history, political regimes have differed greatly from democracies.

This paper shares important similarities with the literature on coalition formation, as analysed by Ray (2007).[4] As in that literature, the process of establishing rules is non-cooperative, but it is assumed that such rules are followed. Moreover, the modelling of rebellions here is related to the idea of blocking in coalitions (Part III of Ray, 2007) in the sense that there is no explicit game-form, as in cooperative game theory. The distinguishing feature of this model is the "rebellion technology" that must be used in order to replace existing institutions by new ones.

The model assumes that the institutions established by the incumbent army determine the allocation of resources once production has taken place. But how would those institutions manage to affect the allocation of goods ex post? As pointed out by Basu (2000) and Mailath, Morris and Postlewaite (2001), laws and institutions do not change the physical nature of the game, all they can do is affect how agents coordinate on some pattern of behaviour. But in reality, laws and institutions are seen to have a strong impact on behaviour, and this feature must be present in any model of institutions.

[4] Baron and Ferejohn (1989) analyse bargaining in legislatures using this approach. Levy (2004) studies political parties as coalitions. Recent contributions include Acemoglu, Egorov and Sonin (2008) Piccione and Razin (2009).

The view of this paper is similar to the application of Schelling's (1960) notion of focal points in the organization of society, as put forward by Myerson (2009). The "rules of the game" are self-enforcing as long as society coordinates on punishing whomever deviates from the rules — and whomever deviates from punishing the deviators. For example, if the laws specify how much an individual must pay or receive from another, and if both expect to be harshly punished if they fail to comply (along with any "higher-order" deviators) then the laws will be self enforcing.

Following this, theorizing about institutions is theorizing about (i) how rules (or focal points) are chosen, and (ii) how rules can change. For example, Myerson (2004) explores the idea of justice as a focal point influencing the allocation of resources in society. This paper takes a more cynical view towards our fellow human beings. Here, the individuals in power choose the laws and institutions to maximize their own payoffs, and those institutions can only be destroyed by a rebellion — wiping out the old institutions, and making way for new ones. There is no modelling of the post-production game.

This paper is also related to the literature on social conflict and predation, surveyed by Garfinkel and Skaperdas (2007).[5] It is easy to envisage how conflict could be important in a state of nature: individuals could devote their time to fighting and stealing from others. However, when there are fights, there are deadweight losses. Thus, people would be better off if they could agree on transfers to avoid conflict. This paper presupposes such deals are possible: producers pay taxes to the group in power, which allocates resources according to some predetermined rules. Here, differently from the literature on conflict, individuals fight to be part of the group that sets the rules, not over what has been produced. Moreover, they fight in groups, not as isolated individuals.

There are theoretical models focused on political issues that lead to inefficiencies in protection of property rights. Examples include Glaeser, Scheinkman and Shleifer (2003), Acemoglu (2008), Guriev and Sonin (2009) and Myerson (2010). Here, the possibility of capital expropriation and the consequent need to protect property rights is just a natural consequence of the possibility of investment and the "rebellion technology" that allows institutions to be destroyed and replaced.

Lastly, it is possible to draw an analogy between this paper and models of democracy (Persson and Tabellini, 2000) in the sense that the "election technology" there is replaced by a "rebellion technology" here.

[5]For instance, see Grossman and Kim (1995), Hirshleifer (1995), and also Hafer (2006), Dal Bó and Dal Bó (2010) for some recent contributions.

2 The model of power and distribution

This section presents the analysis of equilibrium institutions in a simple endowment economy. Subsequent sections extend this analysis to richer environments where there is more scope for institutions to affect economic outcomes.

2.1 Environment

There is an area containing a measure-one population of ex-ante identical individuals. An individual will either be *in power* (a member of the *elite*) or a *worker*. *Institutions* (the "rules of the game") determine the composition of the elite and the allocation of resources. Once institutions are created, the rules they lay down are followed by all unless a *rebellion* occurs, destroying the existing institutions and allowing for the creation of new ones. After institutions are formed, there are opportunities for rebellions, and every individual has the opportunity to rebel. There is no technology that allows any individual to commit to take particular actions subsequently.

Workers have access to a production technology that yields an exogenous endowment of q units of a homogeneous good. Those in power cannot produce, but are able to defend the existing institutions from rebellions by imposing a cost on those participating in a rebellion.[6] Some of the elite may participate in a rebellion; the status of those that do not relies on the survival of the current institutions, so these individuals make up the *incumbent army* defending those institutions.

The sequence of events is as shown in Figure 1. A group of individuals takes power and establishes institutions. There are then opportunities for rebellion. If a rebellion occurs, new institutions are established, potentially changing the elite, with these institutions also being subject to subsequent threats of rebellion. When no rebellions occur, workers receive their endowments, the rules laid down by the prevailing institutions are implemented, and payoffs are received.

2.1.1 Establishing institutions

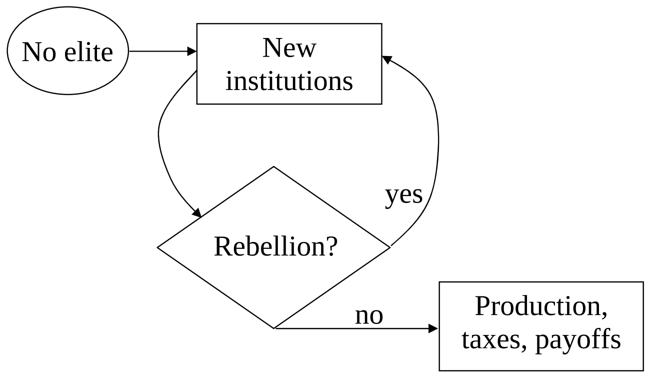
Institutions determine who is in power, and how resources are distributed.

The proportion p of individuals in the elite is such that it maximizes the average utility of those in the elite.[7] In other words, the elite shares power with an extra individual if and only if this increases its average payoff. This assumption captures the idea that the distribution

[6] The assumption that the elite do not receive the same endowment as workers is not essential for the main results. However, it is not unreasonable to suppose there is some opportunity cost for individuals of being in power.

[7] Moving away from the assumption that the elite maximizes its average payoff would require modelling its hierarchy, which is beyond the scope of the paper. See Myerson (2008) for a model addressing this question.

Figure 1: *Sequence of events*



of power reflects the interests of the elite, not the welfare of society. Given the size p, the composition of the elite depends on a predetermined ordering set at the time of the previous rebellion.[8]

The distribution of (lump-sum) taxes $\tau(\cdot)$ levied on workers and the distribution of consumption of members of the elite $C_{\mathrm{p}}(\cdot)$ is set to maximize the average elite payoff. No exogenous restrictions are imposed on taxation beyond the natural bound of what workers possess.[9]

The choices of elite size and the distribution of resources to maximize elite payoffs are constrained by the threat of rebellion, as described below. Once institutions are established, each worker gets to know the tax $\tau(\imath)$ he faces and each member of the elite gets to know his consumption $C_{\mathrm{p}}(\imath)$. Further threats of rebellion are then considered.

2.1.2 A rebellion

Incentives for rebellion depend on individuals' beliefs about what institutions would be created — and hence what payoffs would be received — following a rebellion. Let p'^{e} and $U_{\mathrm{p}}'^{e}$ denote respectively beliefs about the size of the elite and the average utility of its members that will be prescribed by the new institutions if the current ones are destroyed. Let $U(\imath)$ denote the utility of individual $\imath$ if no rebellion takes place so that the current institutions prevail. The difference $U_{\mathrm{p}}'^{e} - U(\imath)$ corresponds to the gain individual $\imath$ obtains if the current institutions are destroyed and he belongs to the subsequent elite. So the maximum cost of fighting effort (in utility units) that an individual $\imath$ who expects to have a place in the subsequent elite is

[8]Since all individuals are ex ante identical, the particular identities of those with whom power is shared have no effect on the elite's average payoff. In a situation where there were no pre-existing institutions, a random ordering is used to determine the composition of the elite.

[9]Taxes can be contingent on any observable variables, but this does not play any role in the endowment-economy version of the model. Taxes cannot be contingent on agents' identities, which would imply that identities become state variables, making the analysis much more complicated. In equilibrium, though, the elite has no incentives to set taxes specifically contingent on identities alone.

willing to make is

$$F(\imath) = U_{\mathrm{p}}^{\prime e} - U(\imath). \qquad [2.1]$$

A rebellion is formally defined as an ordering of the whole population. The corresponding *rebel army* $\mathcal{R}$ is the subset of size p'^e of individuals ordered first. This might include members of the current elite: denote by d the measure of individuals in the rebel army who also belong to the current elite. This rebellion is profitable if:

$$\int_{\mathcal{R}} F(\imath)\mathrm{d}\imath > \delta(p - d). \qquad [2.2]$$

The right-hand side is the total fighting effort required to destroy the current institutions. The rebels have to fight against the incumbent army, which comprises those in power who do not belong to the rebel army. The fighting effort is proportional to the measure of the incumbent army $p - d$. The parameter δ measures the power of the incumbent army. It is assumed that individuals' utility is linear in the amount of fighting effort exerted. Furthermore, total fighting effort is the integral of the effort put in by all individuals in the rebel army $\mathcal{R}$, which is the left-hand side of the inequality above.

Thus for a rebellion to be viable the total fighting effort required to displace the current institutions must be smaller than the total effort those in the rebel army would be willing to make. This quantity of effort depends on beliefs about the subsequent institutions, and these beliefs must be consistent with the insitutions that will actually emerge if the rebellion succeeds. Note that if all members of the elite defect ($d = p$), there is no cost whatsoever of destroying the current institutions.

This approach to modelling the threat of conflict allows for a simple representation of the constraints faced by those choosing institutions, without accounting explicitly for the punches and sword thrusts. By assumption, all the fighting is between those in the incumbent army (drawn only from the elite) and those in the rebel army, which comprises only those individuals who expect a place in the subsequent elite. This means that those who will be workers regardless of the success of a particular rebellion would not fight. The rationale for this assumption is that armies face a problem of incentivizing their members to fight. This incentive problem is likely to be much more severe for those who will not have a place in the elite once the fighting is over. The basic problem is that those designing the new institutions after a rebellion have no interest in honouring past inducements to fight to the extent that these would reduce their own payoffs. Promises to pay workers for fighting (the carrot) are costly by definition, while punishments (the stick) would also be costly because they create disgruntlement with the new institutions and increase incentives for another rebellion. But denying a place in the incumbent army to those who do not exert enough effort does not need to be costly owing to the presence of a pool of identical replacements who would like to join.[10]

2.1.3 Production, taxation and payoffs

Once there is no profitable rebellion against the existing institutions, production takes place, taxes are collected, and payoffs are received.

Let the set of individuals in power be $\mathcal{P}$ and the set of workers be $\mathcal{W}$. Worker $\imath$ subject to institutions imposing a tax $\tau(\imath)$ on him receives consumption

$$C_{\mathrm{w}}(\tau(\imath)) = q - \tau(\imath). \tag{2.3}$$

Total consumption of those in power must be equal to the amount collected in taxes:

$$\int_{\mathcal{P}} C_{\mathrm{p}}(\imath) d\imath = \int_{\mathcal{W}} \tau(\imath) d\imath. \tag{2.4}$$

An individual who receives consumption C obtains utility $U = u(C)$, net of any past fighting costs that have been incurred, and where $u(C)$ is a strictly increasing and weakly concave function.

2.2 Equilibrium

Equilibrium institutions are those that maximize the average utility of those in power subject to there being no profitable rebellion. The choice variables are the size of the elite p, the payoffs of each member of the elite, and a tax distribution $\tau(\cdot)$ specifying a lump-sum tax for each worker $\imath \in \mathcal{W}$. The equilibrium institutions are the result of the following constrained maximization problem:

$$\max_{p,\tau(\cdot)} \frac{1}{p} \int_{\mathcal{P}} U_{\mathrm{p}}(\imath) d\imath \;\; \text{s.t.} \;\; \int_{\mathcal{R}\cap\mathcal{W}} \left(U_{\mathrm{p}}^{\prime e} - U_{\mathrm{w}}(\tau(\imath))\right) d\imath + \int_{\mathcal{R}\cap\mathcal{P}} \left(U_{\mathrm{p}}^{\prime e} - U_{\mathrm{p}}(\imath)\right) d\imath \leq \delta(p-d), \tag{2.5}$$

for all sets $\mathcal{R}$ such that $\mathbb{P}[\mathcal{R}] = p^{\prime e}$, where $p^{\prime e}$ is the belief about the size of the elite if a rebellion succeeds, $U_{\mathrm{p}}^{\prime e}$ is the expected utility of those in the elite under the institutions that would be established in case of a successful rebellion, and $d = \mathbb{P}[\mathcal{R} \cap \mathcal{P}]$ is the measure of defectors from the current elite. The solution to [2.5] depends on $p^{\prime e}$ and $U_{\mathrm{p}}^{\prime e}$.

Once the current institutions are destroyed by a rebellion, new ones are formed to maximize the average utility of those who will be in the new elite. The choice of p' is not constrained by the size of the rebel army $\mathbb{P}[\mathcal{R}]$, the assumption being that the extension of power to an

[10] The formulation necessarily abstracts from coordination issues, and it might be thought that these would be a more pressing problem for the rebels than the incumbent. If this were the case, the power parameter δ could be adjusted to account for such differences.

additional individual must be in the interests of all other members of the elite. The order in which individuals join the elite is set down by the predetermined ordering that characterizes the rebellion. The new elite also maximizes over a tax distribution $\tau'(\cdot)$. In equilibrium, the earlier beliefs p'^e and U'^e_p must be equal to the actual values of p' and U'_p given the absence of uncertainty.

The maximization problem characterizing p' and $\tau'(\cdot)$ is of an identical form to that in [2.5] for p and $\tau(\cdot)$ owing to the irrelevance of history: the cost of conflict is sunk and additive; and the choice of new institutions is not constrained by the size of the rebel army. Thus, there are no state variables, and hence no fundamental reason why different institutions would be chosen at each point. It is therefore natural to focus on Markovian equilibria.

A Markovian equilibrium is a solution to the maximization problem [2.5] with $p = p' = p'^e$; an identical distribution of taxes over workers: $\mathbb{P}[\tau(\imath) \leq \uptau] = \mathbb{P}[\tau'(\imath) \leq \uptau]$ for all $\uptau$; an identical distribution of consumption over members of the elite: $\mathbb{P}[C_\mathrm{p}(\imath) \leq C] = \mathbb{P}[C'_\mathrm{p}(\imath) \leq C]$ for all C. The following result demonstrates some features of any Markovian equilibrium.

Proposition 1 *Any Markovian equilibrium must have the following properties:*

(i) Equalization of workers' payoffs: $U_\mathrm{w}(\imath) = U_\mathrm{w}$ *for all* $\imath$ *(with measure one)*

(ii) Sharing power implies sharing rents: $U_\mathrm{p}(\imath) = U_\mathrm{p}$ *for all* $\imath$ *(with measure one)*

(iii) The set of constraints in [2.5] *is equivalent to a single "no-rebellion" constraint:*

$$U_\mathrm{w}(\tau) \geq U_{\mathrm{p}'} - \updelta \frac{p}{p'}. \qquad [2.6]$$

(iv) Power determines rents: $U_\mathrm{p} - U_\mathrm{w} = \updelta$

PROOF See appendix A.1. ■

These results do not rely on risk aversion (they also hold for a linear utility function). When all workers receive the same endowment, payoff equalization is equivalent to tax equalization. The intuition for the equalization of worker payoffs is that as only a subset of individuals takes part in a rebellion, the elite wants to maximize the utility of the subset with the minimum utility. This is achieved by equalizing workers' utility. Introducing inequalities in payoffs reduces the average utility of a subset of size p' with minimum utility, so it necessarily makes it harder to avoid a rebellion.[11]

An analogous argument implies that heterogeneity in elite payoffs is undesirable because it makes it harder to avoid rebellions from within the elite at the same time as ensure an overall

[11]This result is different from those found in some models of electoral competition such as Myerson (1993). In the equilibrium of that model, politicians offer different payoffs to different agents. But there is a similarity with the model here because in neither case will agents' payoffs depend on their initial endowments.

high payoff for them. Because the cost of rebelling is proportional to the measure of members of the elite who do not defect, any inequality will lead the least satisfied individuals in the elite to participate in a rebellion.

The single "no-rebellion" constraint in the third part of the proposition is the effective constraint faced by the elite when the rebellion includes no defectors and payoffs of workers are equalized. Once this is satisfied, all other constraints are redundant. An immediate consequence of this is that the power parameter δ determines the size of the rents received by members of the elite in a Markovian equilibrium.

Therefore, the maximization problem characterizing the equilibrium institutions has the following recursive form

$$\max_{p,\tau} U_{\mathrm{p}}(p,\tau) \;\text{ s.t. }\; U_{\mathrm{p}}(p',\tau') - \delta\frac{p}{p'} \leq U_{\mathrm{w}}(\tau), \qquad [2.7]$$

where p' and τ' solve an identical problem taking p'' and τ'' as given, and so on.

As tax revenue $(1-p)\tau$ is equally shared among the elite, utility is

$$U_{\mathrm{p}}(p,\tau) = u\left(\frac{1-p}{p}\tau\right).$$

The payoff of those in the elite is increasing in the tax τ and decreasing in the size of the elite p. The tradeoff between the two is represented by the convex indifference curves in Figure 2. The elite has two margins to ensure that it avoids rebellions. It can reduce taxes (the "carrot"), or increase its size, as this means that rebels must fight a larger incumbent army (the "stick"). This corresponds to the upward-sloping no-rebellion constraint. The maximum is at the tangency point. With linear preferences, the constraint is a straight line, as depicted in Figure 2. With risk aversion, the constraint implies that τ would be a concave function of p.

2.3 Examples

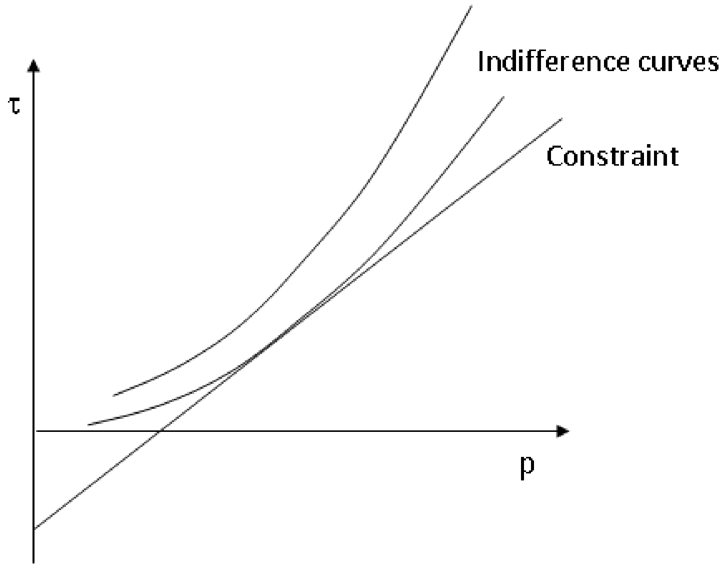
There are two exogenous parameters in the model: the power parameter δ, and the endowment q of a worker. The following examples illustrate the workings of the model for some particular utility functions.

2.3.1 Linear utility

When the utility function is $u(C) = C$, the maximization problem of the elite is

$$\max_{p,\tau} \frac{(1-p)\tau}{p} \;\text{ s.t. }\; C'_{\mathrm{p}} - \delta\frac{p}{p'} \leq q - \tau. \qquad [2.8]$$

Figure 2: *Trade-off between elite size and taxation*



Substituting tax τ from the constraint into the objective function yields:

$$C_{\mathrm{p}} = \frac{1-p}{p}\left(q - C'_{\mathrm{p}} + \delta\frac{p}{p'}\right). \tag{2.9}$$

The following first-order condition with respect to p is obtained:

$$\frac{C_{\mathrm{p}}}{1-p} = (1-p)\frac{\delta}{p'}. \tag{2.10}$$

Imposing Markovian equilibrium ($p = p'$ and $C'_{\mathrm{p}} = C_{\mathrm{p}}$) in [2.9] yields:

$$C_{\mathrm{p}} = (1-p)(q+\delta).$$

Combining this equation with [2.10] (imposing $p = p'$ again) implies:

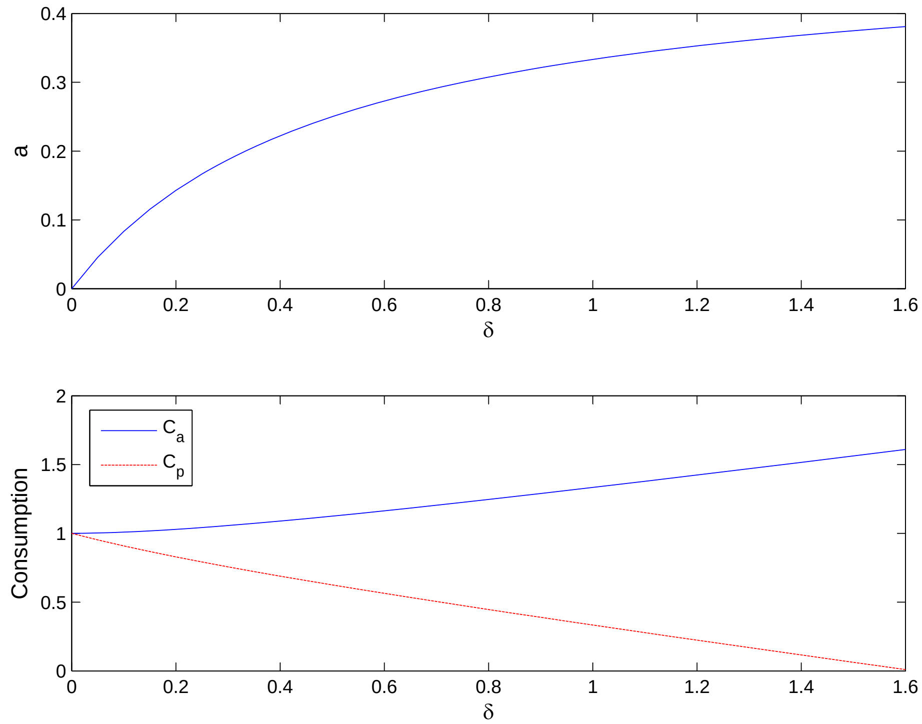
$$p^* = \frac{\delta}{q+2\delta}, \quad C^*_{\mathrm{p}} = \frac{(q+\delta)^2}{q+2\delta}, \quad \text{and } C^*_{\mathrm{w}} = \frac{(q+\delta)^2}{q+2\delta} - \delta.$$

With a linear utility function, the size of the elite is a function of q/δ.

The relationship between the power parameter δ and the key endogenous variables of the model is shown in Figure 3 for $q = 1$.

If $\delta/q \geq (1+\sqrt{5})/2$, consumption of workers is zero. To make the problem interesting, it is sensible to restrict the parameters to ensure workers obtain positive consumption, which requires $\delta/q < (1+\sqrt{5})/2$.

Figure 3: *The case of linear utility*



The power parameter δ affects the equilibrium in three ways. First, an increase in δ makes the incumbent army stronger because the rebels have to bear a higher cost to defeat it. This leads to an increase in τ and a decrease in p. Second, the payoff that the rebels will receive once in power increases as their position would also be stronger once they have supplanted the current elite, making rebellion more attractive. This effect makes the position of the elite weaker, leading it to decrease τ, and increase p. Third, an increase in δ raises the effectiveness of the marginal fighter in the incumbent army, leading the elite to increase its size in order to extract higher taxes. As long as $C_{\mathrm{w}} > 0$, the third effect dominates and the size of the elite is increasing in δ.

2.3.2 Log utility

When the utility function is $u(C) = \log C$, the maximization problem of the elite is

$$\max_{p,\tau} \ \log\left(\frac{(1-p)\tau}{p}\right) \quad \text{s.t.} \quad \log\left(C'_{\mathrm{p}}\right) - \delta\frac{p}{p'} \leq \log(q-\tau). \qquad [2.11]$$

Substituting τ from the constraint into the objective function, the following first-order condition is obtained:

$$\frac{1}{1-p} + \frac{1}{p} = \frac{\delta C'_{\mathrm{p}} \exp\{-\delta p/p'\}}{p'\tau}.$$

Imposing Markovian equilibrium ($p = p'$ and $C'_{\mathrm{p}} = (1-p)\tau/p$) in the above yields:

$$p = (1-p)^2 \delta \exp\{-\delta\},$$

which implies that:

$$p^* = \frac{2\delta \exp\{-\delta\}}{1 + 2\delta \exp\{-\delta\} + \sqrt{1 + 4\delta \exp\{-\delta\}}}.$$

The size of the elite is independent of a worker's endowment q. As consumption of a worker increases, so does the amount of goods he is willing to surrender in order to avoid conflict.

It is also instructive to note how the output of the economy is distributed among individuals. The following condition is obtained by imposing equilibrium and using the no-rebellion constraint:

$$C^*_{\mathrm{p}} = \frac{(1-p)q}{p + (1-p)\exp\{-\delta\}}, \quad \text{and} \quad C^*_{\mathrm{w}} = \frac{\exp\{-\delta\}(1-p)q}{p + (1-p)\exp\{-\delta\}}. \qquad [2.12]$$

The consumption of members of the elite and workers is proportional to q. The output of the economy $(1-p)q$ is divided among all individuals in the economy, with workers getting a fraction $\exp\{-\delta\}$ of what someone in the elite receives. The value of q (affecting the size of the pie) has no influence on the shares of each individual owing to log utility in consumption.

The relationship between the power parameter δ and the key endogenous variables of the model is shown in Figure 4.

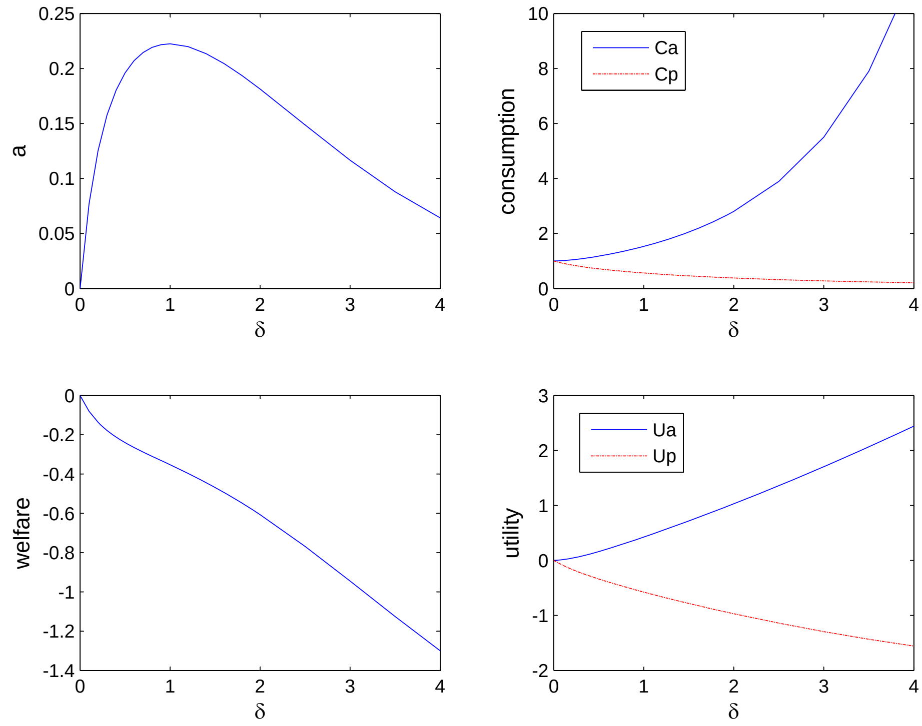
The size of the elite p is positively related to δ for $\delta < 1$ and decreasing in δ otherwise. An increase in δ raises the effectiveness of the marginal fighter in the incumbent army. When δ is relatively small, this leads to an increase in the size of the elite and higher taxes. But as δ becomes larger, reducing the consumption of workers leads to greater incentives for them to rebel, leading the elite to choose a smaller p and a smaller τ.

Output in the economy is negatively related to p. So it reaches its lowest level at $\delta = 1$ when p reaches its maximum. However, welfare is everywhere decreasing in δ owing to the negative distributional effects of increases in δ.

3 The provision of public goods

In the previous section there was no scope for the elite to do what governments are customarily thought to do, such as the provision of public goods. This section introduces a technology

Figure 4: *The case of log utility*



that allows for production of public goods. It is then natural to ask whether such public goods would be provided, since unlikely atomistic individuals, the elite can set up institutions that determine taxes and spending on the provision of public goods. The question is of whether a political Coase Theorem will arise in this setting.

The new technology converts units of output into public goods. If g units of goods per capita are converted using the technology then everyone receives an extra $f(g)$ units of the consumption good. The utility of an individual is thus $u(C + f(g))$.

Per-capita consumption is

$$(1 - p)q - g + f(g).$$

A benevolent social planner would choose g such that

$$f'(g^O) = 1, \qquad [3.1]$$

to maximize the final amount of goods available for consumption. Note that the choice of g is independent of p.

The assumptions of section 2 are now modified so that the institutions include the provision of public goods g. All individuals observe the choice of g and take it into account when determining how much fighting effort they are willing to make in a rebellion.

The consumption of a worker is now

$$C_{\mathrm{w}}(\tau, g) = q - \tau + f(g), \qquad [3.2]$$

and the consumption of a member of the elite is:

$$C_{\mathrm{p}}(p, \tau, g) = \frac{\tau(1-p) - g}{p} + f(g). \qquad [3.3]$$

The equilibrium institutions are the solution of the following maximization problem:

$$\max_{p,\tau,g} u\left(\frac{\tau(1-p) - g}{p} + f(g)\right) \text{ s.t. } U'_{\mathrm{p}} - \delta\frac{p}{p'} \leq U\left(q + f(g) - \tau\right), \\ \text{with } p = p' \text{ , } \tau = \tau' \text{ and } g = g'. \qquad [3.4]$$

Setting up the Lagrangian and taking the first-order conditions with respect to τ and g yields:

$$u'(C_{\mathrm{p}})\left(-\frac{1}{p} + f'(g)\right) = \lambda U'(C_{\mathrm{w}})f'(g),$$
$$u'(C_{\mathrm{p}})\left(\frac{1-p}{p}\right) = -\lambda U'(C_{\mathrm{w}}),$$

and combining both leads to:

$$f'(g^*) = 1.$$

This is the same equation as [3.1] for the case of the benevolent social planner, so the public good is optimally provided.

Who benefits from public good provision? The distribution of output among the individuals depends on the particular utility function. For example, with log utility, the following payoffs analogous to [2.12] are obtained:

$$C^*_{\mathrm{p}} = \frac{(1-p)q + f(g) - g}{p + (1-p)\exp\{-\delta\}}, \quad \text{and } C^*_{\mathrm{w}} = \frac{\exp\{-\delta\}((1-p)q + f(g) - g)}{p + (1-p)\exp\{-\delta\}}.$$

Thus, even though the elite is extracting rents from workers, this does not preclude it from acting as if it were benevolent in other contexts. An implication is that the welfare of workers could be larger or smaller compared to a world in which no-one can compel others to act against their will. This reflects the ambivalence effects of having a ruling elite (or a king) on ordinary people.[12]

The "no-rebellion" constraint implies that the elite cannot disregard the interests of the

[12] As stressed by the prophet Samuel.

workers. Provision of public goods slackens the "no-rebellion" constraint, while the taxes raised to finance them tighten the constraint. By optimally trading off the benefits of the public good against the cost of provision, the elite effectively maximizes the size of the pie, making use of transfers to ensure everyone is indifferent between rebelling or not. By not rebelling, those outside the elite essentially acquiesce to the "offer" made by the elite, analogous to the contracting that underlies the regular Coase Theorem.

The result is far from surprising and can be obtained in several other settings. This is discussed by Persson and Tabellini (2000) in the context of voting and elections. Here the result provides a benchmark where a political Coase Theorem holds.

4 Investment

This section adds the possibility of investment to the model. Workers can exert effort to obtain a larger amount of goods in the future, but there is a time lag between the effort being made and the fruits of that investment being realized. During this span of time, there are opportunities for rebellion against the prevailing institutions. The model is otherwise identical.

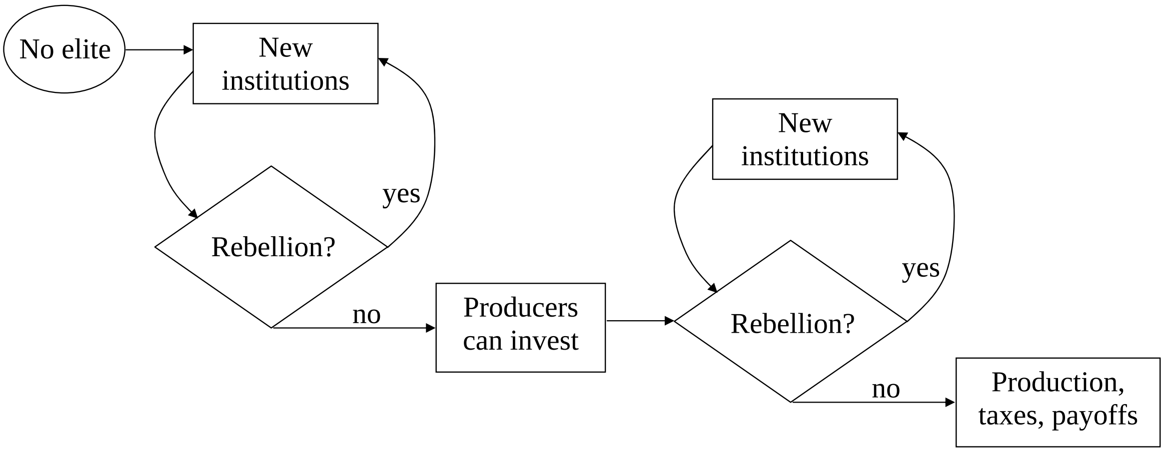
There are no changes here to the mechanism through which institutions are created and destroyed. However, the presence of investment proceeds changes incentives for rebellion, and thus the incentives of the elite over the choice of institutions. The first question is what the equilibrium institutions are. The second is whether those institutions are consistent with a political Coase Theorem.

4.1 Environment

Figure 5 shows the sequence of events in the model. The first part of the sequence resembles the model of section 2 (Figure 1), but now after institutions triggering no rebellion are chosen, some workers receive an investment opportunity and decide whether to invest. Once investment decisions are made, there is another round of opportunities for rebellion, with new institutions established if a rebellion occurs. When the prevailing institutions do not trigger a rebellion, production takes place, taxes are paid, and payoffs are received. In equilibrium, the elite at the pre-investment stage will choose institutions that survive rebellion at all points (as in the model of section 2).

The investment opportunity is as follows. A measure ϕ of individuals gets to know an idiosyncratic effort cost θ (in utility terms), an independent draw from a uniform distribution $U[\theta_L, \theta_H]$. The remaining individuals (measure $1 - \phi$) receive no investment opportunity. Any worker choosing to incur the effort cost receives an extra κ units of goods, referred to as *capital*, at the production and payoffs stage of Figure 5.

Figure 5: *Sequence of events*



It is assumed that the measure of people who get an investment opportunity ϕ is smaller than a threshold:

$$\phi < ..., \qquad [4.1]$$

which ensures that a non-negligible measure of workers will not invest.[13]

$k \in \{0, \kappa\}$ denotes an individual worker's holdings of capital at the post-investment stage. Workers' effort costs θ are private information, while their capital k is publicly observable. This means that taxes can be contingent on capital, but not on effort. As before, lump-sum taxes are available.

For analytical tractability, agents' preferences are linear in consumption, so $u(C) = C$. This allows for relatively simple closed-form solutions.

4.2 Equilibrium

Characterizing the equilibrium requires working backwards from the post-investment stage, determining the equilibrium institutions if a rebellion were to occur, and then analysing what institutions will be chosen at the pre-investment stage.

4.2.1 Post-investment stage after a rebellion

Suppose that a rebellion occurs at some point during the post-investment stage. An argument similar to Proposition 1 shows that the institutions chosen at this point would equalize consumption of all workers, so consumption cannot be contingent on how much capital one

[13] If the parameter restriction in [4.1] does not hold, the nature of the binding constraints might change and the problem becomes more algebrically convoluted. While this analysis could in principle add some twists to the results, it would not affect any of the conclusions in this paper, so it is left for future research.

worker has relative to what others have. The workers' effort costs at the pre-investment stage are sunk, so they do not affect the willingness of invididuals to participate in rebellions.

Let $\hat{U}_{\mathrm{w}}$ denote the utility common to all workers after a rebellion has occurred at the post-investment stage, net of any effort costs (for fighting or investment) incurred earlier. Similarly, $\hat{U}_{\mathrm{p}}$ is the utility of those in power after a post-investment-stage rebellion, and $\hat{p}$ is the size of such an elite. Let K denote the total amount of capital resulting from workers' earlier investment decisions. The economy's resource constraint is

$$\hat{p}\hat{U}_{\mathrm{p}} + (1-\hat{p})\hat{U}_{\mathrm{w}} = K + (1-\hat{p})q,$$

and thus the payoff of those in power is

$$\hat{U}_{\mathrm{p}} = \frac{K + (1-\hat{p})(q-\hat{U}_{\mathrm{w}})}{\hat{p}}.$$

The no-rebellion constraint faced by the elite is

$$\hat{U}_{\mathrm{w}} \leq \hat{U}'_{\mathrm{p}} - \delta\frac{\hat{p}}{\hat{p}'}.$$

Maximizing the elite payoff subject to the no-rebellion constraint and imposing Markovian equilibrium yields:

$$\hat{p} = \frac{\delta}{q+2\delta}, \quad \hat{U}_{\mathrm{p}} = \frac{(q+\delta)^2}{q+2\delta} + K, \quad \text{and } \hat{U}_{\mathrm{w}} = \frac{(q+\delta)^2}{q+2\delta} - \delta + K, \tag{4.2}$$

assuming that δ/q is smaller than $(1+\sqrt{5})/2$ (otherwise $\hat{U}_{\mathrm{w}}$ would be zero).

The size of the elite $\hat{p}$ is the same as that found in section 2.3.1. It is independent of K, which is analytically convenient. Thus were a rebellion to occur at the post-investment stage, the total capital stock K would be expropriated and equally distributed among the whole population.

4.2.2 Equilibrium institutions

The now familiar argument of Proposition 1 implies that the elite has incentives to avoid creating dispersion in workers' payoffs orthogonal to capital holdings. This means that the elite's maximization problem reduces to the choice of its size p, a level of lump-sum taxation τ_q levied on all workers, and a tax τ_κ paid by all those workers who hold κ units of capital.

The choice of capital taxation τ_κ determines workers' effort threshold for investment. A worker will invest if his idiosyncratic effort cost θ is smaller than θ^*, where

$$\theta^* = \kappa - \tau_\kappa. \tag{4.3}$$

This determines the fraction s of workers who will choose to invest:

$$s = \phi\left(\frac{\theta^* - \theta_L}{\theta_H - \theta_L}\right), \tag{4.4}$$

and the capital stock at the second stage, $K = s(1-p)\kappa$.

The payoff of a member of the elite is

$$U_p = \frac{(1-p)\left(\tau_q + s\tau_\kappa\right)}{p}.$$

At the point where the pre-investment institutions are determined, individual workers do not know whether they will receive and investment opportunity and do not know their idiosyncratic effort costs θ. The expected utility of a worker net of effort costs is

$$\mathbb{E}U_w = q - \tau_q + s\mathbb{E}\left[\kappa - \tau_\kappa - \theta | \theta \le \theta^*\right] = q - \tau_q + s\frac{\theta^* - \theta_L}{2}.$$

Therefore, similar to what is obtained in section 2, the first-stage no-rebellion constraint is:

$$q - \tau_q + s\frac{\theta^* - \theta_L}{2} \ge U_p' - \delta\frac{p}{p'} \tag{4.5}$$

The consumption of those workers who choose not to invest is $q - \tau_q$, whereas those who did invest are able to consume $q - \tau_q + \kappa - \tau_\kappa$, which is larger than the former (because $\kappa - \tau_\kappa > 0$). Therefore, workers who have invested are not willing to make as much fighting effort to destroy the current institutions. Consequently, their no-rebellion constraint is not binding, creating incentives for the elite to expropriate their capital.

The no rebellion constraint in the post-investment stage is then:

$$\hat{p}\left(\frac{(q+\delta)^2}{q+2\delta} + (1-p)s\kappa\right) - (\hat{p} - d)(q - \tau_q) - dU_p \le \delta(p - d). \tag{4.6}$$

Any rebellion will comprise d members of the elite and $\hat{p} - d$ workers. The measure of defectors d cannot be larger than p, and the measure of workers in a rebellion $(\hat{p} - d)$ cannot be larger than the measure of workers without capital, $(1-p)(1-s)$. The constraint [4.6] must hold for all values of d in this region.

The following proposition describes some key features of any equilibrium with investment.

Proposition 2 *Any Markovian equilibrium with $s > 0$ must have the following features:*

(i) The workers' pre-investment-stage constraint [4.5] is slack. The following no-rebellion constraints for members of the elite and workers are binding at the post-investment

stage:

$$q - \tau_q \geq \frac{(q+\delta)^2}{q+2\delta} + (1-p)s\kappa - \delta\frac{p}{\hat{p}}, \quad \textit{and} \qquad [4.7]$$

$$U_{\mathrm{p}} \geq \frac{(q+\delta)^2}{q+2\delta} + (1-p)s\kappa - \delta\frac{p}{\hat{p}} + \delta. \qquad [4.8]$$

(ii) The equilibrium size of the elite is

$$p^* = \frac{\delta\hat{p} + s\theta^*}{\delta + s\theta^*}. \qquad [4.9]$$

(iii) The payoff of a member of the elite is

$$U_{\mathrm{p}}^* = \frac{\delta(q+\delta)}{q+2\delta}\left(\frac{q+\delta+s(\kappa-\theta^*)}{\delta+s\theta^*}\right). \qquad [4.10]$$

PROOF See appendix A.2. ■

The constraints [4.7] and [4.8] can be obtained by substituting $d = 0$ and $d = \hat{p}$, respectively, in equation [4.6]. The proposition shows that those constraints bind even when there are not enough workers without capital to launch a rebellion on their own, as long as the constraint on parameters given by [4.1] holds. Owing to the linearity of [4.6] in d, once [4.7] and [4.8] hold, [4.6] must necessarily hold as well.

The binding constraint [4.7] yields a lower bound on the payoff of workers without capital. Constraint [4.8] provides a lower bound to the payoff of those in power, which is exactly what is maximized in this problem. Nonetheless, this constraint must be binding.

To sustain investment, the elite must convince workers that there will be no rebellions leading to expropriation of their capital. The proposition shows that this is equivalent to removing incentives for rebellion from members of the elite and those workers with no capital. The presence of capital increases incentives for rebellion. Thus protection of investment requires an increase in the cost of a rebellion, which can only be achieved by a larger p. Taxes and transfer can only redistribute disgruntlement with the current institutions. On the one hand, discouraging rebellion by workers requires either a larger elite or *lower* taxes, on the other hand, discouraging rebellion by members of the elite requires either a larger elite or *higher* taxes. A larger elite is thus needed to satisfy both constraints simultaneously.

If some workers are to invest then the size of the elite p must be larger than what would prevail if a rebellion were to occur after investment decisions had been made. Then, following a post-investment rebellion, the post-rebellion elite would never include all members of the existing elite. Since some of them would lose their status, there can be no rebellion which commands the unanimous support of the current elite. In the language of the model, for any possible rebellion, even if the rebel army were exclusively drawn from the current elite, others in the elite would constitute the incumbent army that would defend the current institutions against the rebellion, as destruction of their institutions would lead to them becoming workers. This incumbent army would impose costs on those wanting to destroy the current institutions. This cost however would have been zero if all existing members could be guaranteed a place in the post-rebellion elite.

Creating incentives for investment thus requires power sharing: the elite is larger than what would otherwise prevail without the possibility of investment. But importantly, in the model, power is not shared with those who are actually investing. The extra members of the elite have no special function and have no access to any technology directly protecting property rights. Their role is to oppose changes to the status quo, with them fighting against rebellions from workers, and especially from other members of the elite. They do this merely because they would lose from rebellions that dislodge them from the ruling elite.

The choice of τ_κ (which determines s) represents limitations on expropriation, and p captures the distribution of power in society. Credible limitations on expropriation require greater protection of the institutions. Institutions can only be protected from those who hold power if some of them fear losing power if the institutions are destroyed. That can only be achieved if power is not as concentrated as those in power would like it to be, ex-post.

This analysis relies on two key assumptions. First, institutions survive and cannot be modified without a rebellion. This assumption is necessary for the sheer existence of an environment where rules are followed. It captures the idea that elites can in principle create institutions that tie their hands when they would want to tax all the workers' output, expropriate capital or deviate from commitments to provide public goods. If institutions could be easily modified, the "game" would have no "rules". Second, were a rebellion to occur, there can be no enforcement of deals made prior to the rebellion. This reflects the commitment problem highlighted by Acemoglu (2003). The underlying idea is that institutions are needed to enforce deals, but there are no "meta-institutions" to enforce deals concerning the choice of institutions themselves.

Because the destruction of institutions leads to unconstrained optimization over all dimensions of the new institutions, the current elite might struggle to agree to launch a rebellion, even if they are the only participants in the rebellion. If it were possible for optimization following a rebellion to be restricted to certain areas, this might make credible commitments impossible. For example, if the composition of the elite were defined by page one of the constitution, but limitations on expropriation were on page two, then being able to rebel against page two but at the same time commiting not to touch page one would annihilate the credibility of page two.

Differentiating the expression in [4.10] with respect to s, noting that θ^* is a function of s

defined in [4.4], yields the first-order condition:

$$\kappa s^2 + 2(q + 2\delta)s - \frac{\phi}{\theta_H - \theta_L}\left(\delta\kappa - \theta_L(q + 2\delta)\right) = 0$$

where θ^* is also a function of s. One root of this equation is negative. The positive root yields the value for s^* as long as it is between 0 and ϕ, in which case:

$$s^* = \sqrt{\left(\frac{q + 2\delta}{k}\right)^2 + \frac{\phi}{\theta_H - \theta_L}\left(\delta - (q + 2\delta)\frac{\theta_L}{k}\right)} - \frac{q + 2\delta}{k}. \qquad [4.11]$$

Otherwise, the problem yields a corner solution for s. In particular, $s^* = 0$ if

$$\delta\kappa - \theta_L(q + 2\delta) < 0 \Rightarrow \frac{\kappa}{\theta_L} - 1 > 1 + \frac{q}{\delta},$$

and $s^* = \phi$ if

$$\phi < \frac{1}{\theta_H - \theta_L}\left(\delta - (2\theta_H - \theta_L)\frac{q + 2\delta}{\kappa}\right).$$

As mentioned above, the no-rebellion constraint for workers with good investment opportunities is not binding. Thus, the distribution of income does not correspond exactly to the distribution of power. Economic prosperity is related to having more citizens with slack no-rebellion constraints, which requires institutions that protect their investments. The model allows the elite (costlessly) to build those institutions. As in equilibrium agents will only invest if those institutions are there, is in the best interest of the elite to build them?

4.3 The efficient choice of capital taxes

As long as $\kappa > \theta_H$, if capital could be directly protected from expropriation at no cost, investing would always be efficient. However, in this model, that might not be true. If there is investment, more people have to be in the elite in order to prevent a rebellion, which entails an opportunity cost.

The efficient benchmark in this case is the following: taxes on capital τ_κ are chosen by a benevolent agent, knowing that everything else (p and τ_q) will be determined in equilibrium, from the problem of maximizing payoffs of the group in power subject to the threat of rebellions, as discussed above.

In the case of section 3, such a benevolent agent could not create a better economic outcome with a different choice of g. In that case, the public good is efficiently provided. Here, is protection of property rights efficiently provided? If not, is it underprovided or overprovided and why?

For a given τ_κ, restrictions determine p and τ_q. The capital tax τ_κ determines s, which pins down the elite size, p^*, as any smaller p would lead to a rebellion at the second stage. Taxes on q would be such that the constraint in [4.7] is binding.

Welfare in the economy is given by:

$$W = pU_{\text{p}} + (1-p)U_{\text{w}} + (1-p)s\frac{\theta^* - \theta_L}{2}$$

Since $U_{\text{p}} = U_{\text{w}} + \delta$ in equilibrium,

$$W = U_{\text{p}} - (1-p)\delta + (1-p)s\frac{\theta^* - \theta_L}{2} \qquad [4.12]$$

There are two differences between the expressions for W and U_{p}. The second term in [4.12] is related to the distribution of resources in the economy. The third term is the investor surplus.

Substituting $p = p^*$ and rearranging:

$$W = \frac{\delta(q+\delta)}{q+2\delta}\left(\frac{q + s(\kappa - \theta^*)}{\delta + s\theta^*} + \frac{s^2(\theta_H - \theta_L)}{2\phi(\delta + s\theta^*)}\right)$$

The first-order condition with respect to s is obtained by differentiating the above equation, using [4.4]:

$$\frac{2\kappa - \theta_L}{2}s^2 + (\delta + 2q)s - \frac{\phi}{\theta_H - \theta_L}(\delta\kappa - (q+\delta)\theta_L) = 0$$

One root of this equation is negative. The other yields the solution for the optimal fraction of investors s^O, as long as it is between 0 and ϕ. In that case,

$$s^O = \frac{\sqrt{(\delta+2q)^2 + 2\phi\frac{2\kappa-\theta_L}{\theta_H-\theta_L}(\delta\kappa - (q+\delta)\theta_L)} - (\delta+2q)}{2\kappa - \theta_L} \qquad [4.13]$$

The following proposition establishes the breakdown of the political Coase Theorem when private investment is possible but subject to expropriation:

Proposition 3 *Unless s^O and s^* are both equal to 0 or ϕ,*

$$s^O > s^*$$

Proof See appendix A.3. ■

The constrained welfare-maximizing choice of s leads to more investment than the equilibrium choice for two reasons: first, the equilibrium choice of capital taxes does not take into account investors' surpluses (the third term in [4.12]). As θ is not observable, it is impossible for the elite to extract rents from those who invested, at the margin. Consequently, the no-rebellion constraint for those who invested is slack. A political Coase theorem does not hold here because of the non-observability of effort, θ, which makes it impossible to write rules contingent on it.

The second (and more interesting) distortion follows from the distributional effects of protecting against expropriation (the second term in [4.12]). Protection of property rights requires sharing power. Sharing power requires sharing rents, because a rebellion including members of the elite is less costly in terms of fighting effort. Thus while a larger elite might allow for profitable investment and higher output, it also reduces the fraction of output appropriated by a member of the elite. Hence, in equilibruim, institutions leading to lower output might arise owing to the impact of the distribution of power on the distribution of income.

The association between power and rents sets an endogenous limit on the set of possible transfers, leading to a break down of a political Coase Theorem. The elite could implement the constrained-optimal allocation, which requires a larger p, but the threat of rebellion constrains their choice of the extra elite members' payoffs.

The inefficiently high capital taxes can be interpreted as insufficient protection against expropriation of property. Recent empirical work has highlighted the importance of institutions that prevent the government from expropriating individuals' resources. But why would expropriation of property be so susceptible to political failures? Why would a political Coase Theorem fail to apply in this particular case? The model here sheds light on this question.

Last, a note on the effect of the power parameter δ. In section 3, the welfare of workers in an economy with institutions and a self-interested elite might be larger than the welfare of workers in an economy where no transfers can be enforced. The possibility of establishing rules allows for the provision of public goods. However, a larger δ can only harm the workers, as it makes possible higher taxes and leads to a worse income distribution. In contrast, in an economy with a very small δ, there cannot be investment (s is always equal to 0). If δ is low enough, it is not even efficient to protect property – that would require a very large increase in the elite, which entails an opportunity cost. A larger δ allows the elite to keep its power, which will be used for their own good, but might also allow them to (partly) protect private property.

4.4 Analogies with historical examples

The results show the importance of sharing power to stimulate investment and prevent changes in the rules of the game that lead to expropriation. They also show that rulers will not share power as much as would be efficient. Although the model is too abstract to match any given historical example precisely, the results resonate with a number of examples.

Broadly speaking, the extra individuals in power required to protect property rights ($p^* - \hat{p}$) might be interpreted as a "parliament" or any other group of people with power to deter changes in institutions – from outsiders and, especially, from insiders. Parliaments are usually thought of as representing those who elected their members, but so are democractically elected presidents, and members of parliaments are not only useful for defending minorities. Power sharing makes institutions more stable because it makes it costly for some members of the elite to replace the current institutions with new ones – with potentially different rules on how power is distributed and on the limits to how much can be extracted from producers. Once power is concentrated, institutions are subject to the whims of those in power. In the words of Montesquieu, "Princes who have wanted to make themselves despotic have always begun by uniting in their person all magistracies."

Malmendier (2009) studies what he considers the earliest precursors of the modern business corporation, the Roman *societas publicanorum*. Their demise occurred with the transition from the Roman Republic to the Roman Empire. According to Malmendier (2009), one possible reason for their demise is that "the Roman Republic was a system of checks and balances. But the emperors centralized power and could, in principle, bend law and enforcement in their favor". In other words, while power was decentralized, it was possible to have rules that guaranteed the property rights of the *societas publicanorum*, presumably because changing the rules would require some of the individuals in power to launch an attack on some of their peers, which was costly. Once power was centralized, protection agaisnt expropriation was not possible any longer.

Greif (2006) analyses the importance of the *podesteria* system for interclan cooperation in medieval Italy. The *podestà*, an individual coming from another city who would be in power for a year, was generously paid and played an important role in the building of cities such as Genoa by allowing for cooperation and investment. Interestingly, Greif argues that the *podestà* had to be sufficiently strong because otherwise, if one clan had defeated its rival, it could easily defeat the *podestà* as well. Translating this into the language of the model, the strength of the *podestà* ($p^* - \hat{p}$) had to be large enough to ensure he could not be defeated by a clan, which would then be able to change the institutions.[14]

In seventeenth-century Britain, the Glorious Revolution led to power sharing between King and Parliament. By accepting the Bill of Rights, King William III accepted that power would be shared between him and the parliament. North and Weingast (1989) argued that the Glorious Revolution led to secure property rights and elimination of confiscatory government. Shortly after the Glorious Revolution, the government could borrow much more, and at substantially lower rates. This was certainly in the interest of the king, yet the Stuart kings had strongly resisted sharing power with Parliament. According to the model, secure property

[14] Greif (2006), pp. 239–240.

rights require just such power sharing, so that it is costly for the king to rewrite the rules ex post. However, the existence of a parliament with real power implies that rents have to be shared, so even if the pie becomes larger, the share of the king might be smaller.

5 Conclusions

Governments and institutions play a key role in development by making it possible to direct resources to the provision of public goods and by enforcing collective choices. However, what is called a government in some of the poorest countries bears little resemblance to its counterpart in the most developed nations. As noted in the book of Samuel, those with power to choose the "rules of the game", guided by their own interests, might choose policies that are not necessarily good from the point of view of society. Indeed, in far too many cases, the warnings of Samuel remain as relevant as ever.[15]

This paper provides a model where institutions are set to maximize the payoffs of the group in power. That assumption and the threat of rebellions are the main elements of the model. When private investment is considered, a risk of expropriation emerges. In order to protect investment, sharing power is needed. That is the only way to prevent those in power from tearing up the old rules and expropriating capital. However, sharing power requires sharing rents, because those with power can use it against the other members of the elite. This imposes limits to the available set of transfers and leads to the breakdown of a political Coase Theorem. There is not enough power sharing and consequently capital taxes are excessively high — too little protection of investment.

This framework could be used in a number of ways. First, it could be explored to understand which factors allow for more or less power sharing — and thus investment. It would also be interesting to analyse other sources of inefficiencies. Also, agents are homogeneous ex ante, with heterogeneity is an equilibrium outcome, but ex-ante heterogeneity of various forms could be incorporated into the model (for example, ethnic differences, as in Caselli and Coleman (2006)). There are no rebellions in equilibrium, but these could arise if the model had some stochastic elements added.

References

Acemoglu, D. (2003), "Why not a political Coase theorem? Social conflict, commitment, and politics", *Journal of Comparative Economics*, **31(4)**:620–652. 2, 23

——— (2008), "Oligarchic versus democratic societies", *Journal of the European Economic Association*, **6(1)**:1–44 (March). 6

[15] For a model of the trade-off that having a government (or a "king") entails, see Grossman (2002).

Acemoglu, D., Egorov, G. and Sonin, K. (2008), "Coalition formation in non-democracies", *Review of Economic Studies*, **75(4)**:987–1009 (October). 5

Acemoglu, D. and Johnson, S. (2005), "Unbundling institutions", *Journal of Political Economy*, **113(5)**:949–995. 4

Acemoglu, D., Johnson, S. and Robinson, J. A. (2005), "Institutions as a fundamental cause of long-run growth", in P. Aghion and S. N. Durlauf (eds.), *Handbook of Economic Growth*, North-Holland, chap. 6, pp. 386–472. 1, 2

Acemoglu, D. and Robinson, J. A. (2006), *Economic Origins of Dictatorship and Democracy*, Cambridge University Press. 1

——— (2008), "Persistence of power, elites, and institutions", *American Economic Review*, **98(1)**:267–293 (March). 1

Baron, D. P. and Ferejohn, J. A. (1989), "Bargaining in legislatures", *American Political Science Review*, **83(4)**:1181–1206 (December). 5

Basu, K. (2000), *Prelude to Political Economy*, Oxford University Press. 5

Besley, T. and Persson, T. (2009a), "The origins of state capacity: Property rights, taxation, and politics", *American Economic Review*, **99(4)**:1218–1244 (September). 1

——— (2009b), "Repression or civil war?", *American Economic Review*, **99(2)**:292–297 (May). 1

——— (2010), "State capacity, conflict and development", *Econometrica*, forthcoming. 1

Caselli, F. and Coleman, W. J., II (2006), "On the theory of ethnic conflict", Working paper, London School of Economics. 28

Dal Bó, E. and Dal Bó, P. (2010), "Workers, warriors and criminals: Social conflict in general equilibrium", *Journal of the European Economic Association*, forthcoming. 6

Downs, A. (1957), "An economic theory of political action in a democracy", *Journal of Political Economy*, **65(2)**:135–150 (April). 5

Engerman, S. L. and Sokoloff, K. L. (1997), "A view from economic historians of the United States", in S. H. Haber (ed.), *How Latin America fell behind*, Stanford University Press, chap. 10, pp. Haber, Stephen H. 1

Garfinkel, M. R. and Skaperdas, S. (2007), "Economics of conflict: An overview", in T. Sandler and K. Hartley (eds.), *Handbook of Defense Economics*, vol. 2, North-Holland, chap. 22, pp. 649–709. 6

GLAESER, E., SCHEINKMAN, J. AND SHLEIFER, A. (2003), "The injustice of inequality", *Journal of Monetary Economics*, **50(1)**:199–222 (January). 6

GREIF, A. (2006), *Institutions and the Path to the Modern Economy*, Cambridge University Press. 1, 27

GROSSMAN, H. I. (2002), "'Make us a king': anarchy, predation, and the state", *European Journal of Political Economy*, **18(1)**:31–46 (March). 28

GROSSMAN, H. I. AND KIM, M. (1995), "Swords or plowshares? A theory of the security of claims to property", *Journal of Political Economy*, **103(6)**:1275–1288 (December). 6

GURIEV, S. AND SONIN, K. (2009), "Dictators and oligarchs: A dynamic theory of contested property rights", *Journal of Public Economics*, **93(1–2)**:1–13 (February). 6

HAFER, C. (2006), "On the origins of property rights: Conflict and production in the state of nature", *Review of Economic Studies*, **73(1)**:119–143 (February). 6

HALL, R. E. AND JONES, C. I. (1999), "Why do some countries produce so much more output per worker than others?", *Quarterly Journal of Economics*, **114(1)**:83–116 (February). 1

HIRSHLEIFER, J. (1995), "Anarchy and its breakdown", *Journal of Political Economy*, **103(1)**:26–52 (February). 6

LEVY, G. (2004), "A model of political parties", *Journal of Economic Theory*, **115(2)**:250–277 (April). 5

MAILATH, G., MORRIS, S. AND POSTLEWAITE, A. (2001), "Laws and authority", Working paper, Princeton University. 5

MALMENDIER, U. (2009), "Law and finance at the origin", *Journal of Economic Literature*, **47(4)**:1076–1108. 27

MYERSON, R. B. (1993), "Incentives to cultivate favored minorities under alternative electoral rules", *American Political Science Review*, **87(4)**:856–869 (December). 11

——— (2004), "Justice, institutions and multiple equilibria", *Chicago Journal of International Law*, **5(1)**:91–107. 6

——— (2008), "The autocrat's credibility problem and foundations of the constitutional state", *American Political Science Review*, **102(1)**:125–139. 7

——— (2009), "Learning from Schelling's *Strategy of Conflict*", *Journal of Economic Literature*, **47(4)**:1109–1125. 6

——— (2010), "Capitalist investment and political liberalization", *Theoretical Economics*, **5(1)**:73–91. 6

North, D. C. (1990), *Institutions, institutional change, and economic performance*, Cambridge University Press. 1

North, D. C. and Weingast, B. R. (1989), "Constitutions and commitment: The evolution of institutional governing public choice in seventeenth-century England", *Journal of Economic History*, **49(4)**:803–832 (December). 1, 27

Persson, T. and Tabellini, G. (2000), *Political Economics: Explaining Economic Policy*, MIT Press. 5, 6, 18

Piccione, M. and Razin, R. (2009), "Coalition formation under power relations", *Theoretical Economics*, **4(1)**:1–15 (March). 5

Ray, D. (2007), *A Game-Theoretic Perspective on Coalition Formation*, Oxford University Press. 5

Schelling, T. C. (1960), *The Strategy of Conflict*, Harvard University Press. 6

A Technical appendix

A.1 Proof of Proposition 1

Start by fixing a particular choice of institutions, that is, a choice of army size p and the distribution of consumption across agents. These choices also determine the composition of $\mathcal{P}$ and $\mathcal{W}$. This particular choice of institutions does not lead to any rebellion if and only if

$$\max_{\mathcal{R}} \left\{ \int_{\mathcal{R}\cap\mathcal{W}} (U'_{\mathrm{p}} - U_{\mathrm{w}}(\imath))d\imath + \int_{\mathcal{R}\cap\mathcal{P}} (U'_{\mathrm{p}} - U_{\mathrm{p}}(\imath))d\imath + \delta d \right\} \leq \delta p \ \text{ s.t. } \ \mathbb{P}[\mathcal{R}] = p',$$

where the size of the subsequent incumbent army p', as well as U'_{p}, is taken as given. For a given $\mathcal{R}$, d is the measure of the set $\mathcal{R}\cap\mathcal{P}$. Now define disjoint sets $\mathcal{R}_w \equiv \mathcal{R}\cap\mathcal{W}$ and $\mathcal{R}_p \equiv \mathcal{R}\cap\mathcal{P}$, and note that the maximization problem above is equivalent to

$$\max_{\mathcal{R}_w,\mathcal{R}_p} \left\{ (p'-d)U'_{\mathrm{p}} - \int_{\mathcal{R}_w} U_{\mathrm{w}}(\imath)d\imath + d(U'_{\mathrm{p}}+\delta) - \int_{\mathcal{R}_p} U_{\mathrm{p}}(\imath)d\imath \right\} \leq \delta p \ \text{ s.t. } \ \mathbb{P}[\mathcal{R}_{\mathrm{w}}\cup\mathcal{R}_{\mathrm{p}}] = p'. \quad \text{[A.1.1]}$$

Now make the following definitions:

$$\bar{U}_{\mathrm{w}} \equiv \frac{1}{1-p}\int_{\mathcal{W}} U_{\mathrm{w}}(\imath)d\imath, \quad \bar{U}_{\mathrm{p}} \equiv \frac{1}{p}\int_{\mathcal{P}} U_{\mathrm{p}}(\imath)d\imath, \quad \text{and} \quad \text{[A.1.2a]}$$

$$\underline{U}_{\mathrm{w}}(r_{\mathrm{w}}) = \min_{\mathcal{R}_{\mathrm{w}} \ \text{s.t.} \ \mathbb{P}[\mathcal{R}_{\mathrm{w}}]=r_{\mathrm{w}}} \left\{ \frac{1}{r_{\mathrm{w}}}\int_{\mathcal{R}_{\mathrm{w}}} U_{\mathrm{w}}(\imath)d\imath \right\}, \quad \underline{U}_{\mathrm{p}}(r_{\mathrm{p}}) = \min_{\mathcal{R}_{\mathrm{p}} \ \text{s.t.} \ \mathbb{P}[\mathcal{R}_{\mathrm{p}}]=r_{\mathrm{p}}} \left\{ \frac{1}{r_{\mathrm{p}}}\int_{\mathcal{R}_{\mathrm{p}}} U_{\mathrm{p}}(\imath)d\imath \right\}. \quad \text{[A.1.2b]}$$

These definitions imply

$$\underline{U}_{\mathrm{w}}(r_{\mathrm{w}}) \leq \bar{U}_{\mathrm{w}}, \tag{A.1.3a}$$

with equality if and only if (i) $C_{\mathrm{w}}(\imath) = C_{\mathrm{w}}$ for all $\imath$ (with measure one), or (ii) $u''(\cdot) = 0$ and $r_{\mathrm{w}} = 1-p$. Analogously,

$$\underline{U}_{\mathrm{p}}(r_{\mathrm{p}}) \leq \bar{U}_{\mathrm{p}}, \tag{A.1.3b}$$

with equality if and only if (i) $C_{\mathrm{p}}(\imath) = C_{\mathrm{p}}$ for all $\imath$ (with measure one), or (ii) $u''(\cdot) = 0$ and $r_{\mathrm{p}} = p$.

The definitions in [A.1.2a] imply that the constraint [A.1.1] is equivalent to

$$\max_{d \in \mathscr{D}(p,p')} \left\{ (p'-d)U'_{\mathrm{p}} - (p'-d)\underline{U}_{\mathrm{w}}(p'-d) + d(U'_{\mathrm{p}} + \delta) - d\underline{U}_{\mathrm{p}}(d) \right\} \leq \delta p, \tag{A.1.4}$$

where $\mathscr{D}(p,p') \equiv [\max\{0, p+p'-1\}, \min\{p,p'\}]$. Rearranging terms yields

$$p'(U'_{\mathrm{p}} - \bar{U}_{\mathrm{w}}) + \max_{d \in \mathscr{D}(p,p')} \left\{ d(\bar{U}_{\mathrm{w}} - \bar{U}_{\mathrm{p}} + \delta) + (p'-d)(\bar{U}_{\mathrm{w}} - \underline{U}_{\mathrm{w}}(p'-d)) + d(\bar{U}_{\mathrm{p}} - \underline{U}_{\mathrm{p}}(d)) \right\} \leq \delta p, \tag{A.1.5}$$

making use of the definitions in [A.1.2a] again.

Now consider a Markovian equilibrium with $p = p'$ and $U'_{\mathrm{p}} = \bar{U}_{\mathrm{p}}$. Let

$$D \equiv \max_{d \in [\max\{0, 2p-1\}, p]} \left\{ d(\bar{U}_{\mathrm{w}} - \bar{U}_{\mathrm{p}} + \delta) + (p-d)(\bar{U}_{\mathrm{w}} - \underline{U}_{\mathrm{w}}(p-d)) + d(\bar{U}_{\mathrm{p}} - \underline{U}_{\mathrm{p}}(d)) \right\}, \tag{A.1.6}$$

where d^* is the smallest value (without loss of generality) of d that maximizes the above expression. The single constraint [A.1.4] to which the equilibrium institutions are subject must be binding. Thus, by substituting the solution of the problem in [A.1.6] into [A.1.4], and after some rearrangement:

$$p\bar{U}_{\mathrm{p}} - d^*\underline{U}_{\mathrm{p}}(d^*) = (p - d^*)(\delta + \underline{U}_{\mathrm{w}}(p - d^*)). \tag{A.1.7}$$

First, note that $d^* < p$, otherwise higher taxes on workers would be feasible. Substituting the inequalities in [A.1.3a] and [A.1.3b] into [A.1.7] yields

$$\bar{U}_{\mathrm{w}} - \bar{U}_{\mathrm{p}} + \delta \geq 0,$$

with equality if payoffs are equalized among army members and among workers. Thus, if payoffs are equalized, then the value of D in [A.1.6] is necessarily zero.

Now suppose that payoffs among the workers are not equalized. Then $(\bar{U}_{\mathrm{w}} - \underline{U}_{\mathrm{w}}(p-d))$ is strictly positive for $d > 2p-1$ since this ensures $p - d < 1 - p$, the total measure of workers. Moreover, for $d < p$, the coefficient multiplying this term in [A.1.6] is positive as well. As the other terms in [A.1.6] would be non-negative, and since such a choice of d is feasible (as $p < 1$), the value of D would be positive. As has been seen, $D = 0$ is possible if institutions equalize payoffs within groups. This can be achieved without affecting the other terms in the constraint [A.1.5], and without decreasing the objective function. Hence, in a Markovian equilibrium, payoffs of workers must be equalized.

Analogously, if payoffs among army members are not equalized, then $(\bar{U}_{\mathrm{p}} - \underline{U}_{\mathrm{p}}(d))$ is strictly positive for $d < p$. For $d > 0$, this term has a positive coefficient in [A.1.6]. Such a choice is feasible ($p < 1$), and the other terms appearing in [A.1.6] would be non-negative. Again, D would hence be positive. Since $D = 0$ is feasible, a similar argument to that above shows that payoffs of army members must be equalized.

A.2 Proof of Proposition 2

0. If $s > 0$, $p^* > \hat{p}$.

If $s > 0$, consumption at the second stage is not the same for all workers: those who invested have to get more than the others (to compensate for effort). So, the no-rebellion constraint of those workers who invested cannot be binding. Now, that implies that the army would benefit from expropriating their capital, so that their consumption would be exactly the same as the consumption of the other workers. But if $p^* \leq \hat{p}$, a rebellion in the second stage is costless for the members of the incumbent army: they can expropriate capital from those who invested at no cost. But that cannot happen in equilibrium.

1. If $s > 0$, [4.5] and [4.8] cannot both bind.

Suppose they do, then imposing equilibrium, they imply:

$$q - \tau_q + s\frac{\theta^* - \theta_L}{2} = U_{\mathrm{p}} - \delta$$

$$U_{\mathrm{p}} = \frac{(q+\delta)^2}{q+2\delta} + (1-p)s\kappa - \delta\frac{p}{\hat{p}} + \delta$$

Plugging the value of U_{p} in the first one,

$$q - \tau_q + s\frac{\theta^* - \theta_L}{2} = \frac{(q+\delta)^2}{q+2\delta} + (1-p)s\kappa - \delta\frac{p}{\hat{p}}$$

which contradicts [4.7].

2. [4.8] cannot be the only one that binds.

Trivial.

3. [4.5] cannot be the only one that binds.

Suppose is the only one that binds. Getting τ_q from equation [4.5] and substituting it at the objective function, we get:

$$U_{\mathrm{p}} = \frac{1-p}{p}\left(q - U'_{\mathrm{p}} + \delta\frac{p}{p'} + s\frac{2\kappa - \theta^* - \theta_L}{2}\right)$$

The first order condition with respect to p yields

$$p^* = \frac{\delta}{q + 2\delta + s\frac{2\kappa - \theta^* - \theta_L}{2}} < \hat{p}$$

But $p^* < \hat{p}$ is a contradiction.

4. So [4.7] binds. Using the constraint to get an expression for τ and substituting it in the objective function, we obtain

$$U_{\mathrm{p}} = \frac{1-p}{p}\left(ps\kappa + \delta\frac{p}{\hat{p}} - s\theta^* - \frac{\delta^2}{q+2\delta}\right) \qquad \text{[A.2.1]}$$

5. Substituting the expression for U_{p} from [A.2.1] into [4.8] and doing algebra, we get:

$$p^* \geq \frac{\delta\hat{p} + s\theta^*}{\delta + s\theta^*} \qquad \text{[A.2.2]}$$

6. Now need to show that [4.8] is binding. To show that, (a) we show a contradiction when assuming that only [4.7] is binding and (b) we show a contradiction when assuming that [4.5] and [4.7] are binding.

6a. If [4.7] is the only constraint binding, the solution is given by the unconstrained maximization of [A.2.1]. Taking the first order condition with respect to p and rearranging yields

$$p^* = \frac{\delta\hat{p} + s\theta^*}{p^*(s\kappa + q + 2\delta)}$$

Condition [A.2.2] implies

$$\frac{\delta\hat{p} + s\theta^*}{p^*(s\kappa + q + 2\delta)} \geq \frac{\delta\hat{p} + s\theta^*}{\delta + s\theta^*}$$

Thus

$$p^*(s\kappa + q + 2\delta) \leq \delta + s\theta^* \Rightarrow p^* \leq \frac{\delta + s\theta^*}{s\kappa + q + 2\delta}$$

Using again condition [A.2.2],

$$\frac{\delta + s\theta^*}{s\kappa + q + 2\delta} \geq \frac{\delta\hat{p} + s\theta^*}{\delta + s\theta^*}$$

Rearranging shows a contradiction.

6b. The Lagrangian of the problem of maximizing U_{p} subject to [4.5] and [4.7] is:

$$\mathcal{L} = \frac{1-p}{p}\left(\tau_q + s(\kappa - \theta^*)\right) + \aleph_1\left(q - \tau_q + s\frac{\theta^* - \theta_L}{2} - U_{\mathrm{p}}' + \delta\frac{p}{p'}\right) + \aleph_2\left(-\tau_q - \frac{\delta^2}{q + 2\delta} - (1-p)s\kappa + \delta\frac{p}{\hat{p}}\right)$$

The first order condition with respect to τ_q implies

$$\aleph_2 = \frac{1-p}{p} - \aleph_1$$

The first order condition with respect to p implies:

$$-\frac{1}{p^2}\left(\tau_q + s(\kappa - \theta^*)\right) + \aleph_1\frac{\delta}{p'} + \aleph_2\left(s\kappa + \frac{\delta}{\hat{p}}\right) = 0$$

Substituting the value of τ_q implied by the binding constraint [4.7] and the value of $\aleph_2$ from the first order with respect to τ_q and rearranging yields

$$\frac{1}{p^2}\left(\delta\hat{p} + s\theta^*\right) = s\kappa + q + 2\delta + \aleph_1\left(s\kappa + \frac{\delta}{\hat{p}} - \frac{\delta}{p}\right)$$

As in equilibrium $p \geq \hat{p}$ and $\aleph_1 \geq 0$, the last term is positive. Hence

$$\frac{1}{p^2}\left(\delta\hat{p} + s\theta^*\right) > s\kappa + q + 2\delta \Rightarrow \frac{\delta\hat{p} + s\theta^*}{p^*(s\kappa + q + 2\delta)} > p^*$$

Now using the same steps as 6a, we get the proof.

As [A.2.2] is a translation of [4.8] using the objective function and the binding constraint [4.7], it is binding and yields the equilibrium value of p^*.

Rearranging [A.2.1] yields

$$U_{\mathrm{p}} = (1-p)(s\kappa + q + 2\delta) - \frac{1-p}{p}\left(\frac{\delta^2}{q + 2\delta} + s\theta^*\right),$$

and substituting the value of p implies equation [4.10].

A.3 Proof of Proposition 3

The value of s^* is given by the positive root of the quadratic:

$$\Sigma^*(s) = \kappa s^2 + 2(q + 2\delta)s - \frac{\phi}{\theta_H - \theta_L}(\delta\kappa - (q + 2\delta)\theta_L)$$

and the value of s^O is given by the positive root of:

$$\Sigma^O(s) = \frac{2\kappa - \theta_L}{2}s^2 + (\delta + 2q)s - \frac{\phi}{\theta_H - \theta_L}(\delta\kappa - (q + \delta)\theta_L)$$

Now,

$$\Sigma^*(s) - \Sigma^*(O) = \frac{\theta_L}{2}s^2 + 3\delta s + \frac{\phi\delta\theta_L}{\theta_H - \theta_L}$$

which is positive for all $s > 0$. As both quadratic functions are convex and have at least one negative root, any positive root of $\Sigma^*(s)$ must be smaller than a positive root of $\Sigma^O(s)$.