REAP Rede de Economia Aplicada

Power Sharing, Rents, and Commitment

Bernardo Guimarães
Kevin D. Sheedy

Julho, 2013
Working Paper 59

POWER SHARING, RENTS, AND COMMITMENT

Bernardo Guimarães
Kevin D. Sheedy

Bernardo Guimarães
Escola de Economia de São Paulo
Fundação Getúlio Vargas (EESP/FGV)
Rua Itapeva, nº 474
01332-000 - São Paulo, SP - Brasil
bernardo.guimaraes@fgv.br

Kevin Sheedy
London School of Economics and Political Science (LSE)
Houghton Street
London WC2A 2AE
Reino Unido
k.d.sheedy@lse.ac.uk

Power Sharing, Rents, and Commitment*

Bernardo Guimaraes[†]
Sao Paulo School of Economics–FGV

Kevin D. Sheedy
London School of Economics

July 2013

Abstract

Good government requires some commitment to rules. This paper proposes a model of the equilibrium rules allocating power and resources that emerge under the threat of rebellions. Commitment to protect property rights can only be achieved if power is not as concentrated as incumbents would like it to be, ex post. Power sharing endogenously enables incumbents to commit to otherwise time-inconsistent policies by ensuring more people receive rents under the status quo, and thus want to defend it. But it is precisely because sharing power entails sharing rents that power is too concentrated in equilibrium, leading to inefficiently low investment.

JEL CLASSIFICATIONS: D74; H11; P48.

KEYWORDS: power sharing; commitment; power struggle; property rights; time inconsistency.

*We thank Bruno Decreuse, Erik Eyster, Emmanuel Farhi, Carlos Eduardo Goncalves, Ethan Ilzetzki, Peter Jensen, Per Krusell, Dirk Niepelt, Andrea Prat, Ronny Razin, Alwyn Young, and seminar participants at U. Amsterdam, the Anglo-French-Italian Macroeconomic workshop, U. Bonn, U. Carlos III, Central European U., Columbia U., CREI conference 'The political economy of economic development', U. Cyprus, EEA 2011, EIEF, U. Erlangen-Nuremberg, ESSIM 2010, FGV–Rio, U. Glasgow, Harvard U., U. Helsinki, HSE-ICEF Moscow, ISI–Delhi Economic Growth and Development conference 2010, Institute for International Economic Studies, the Joint French Macroeconomic Workshop, London School of Economics, U. Oslo, Paris School of Economics, PUC–Rio, RES 2011, U. Sao Paulo, Sao Paulo School of Economics–FGV, SED 2011, U. Surrey, U. St. Gallen, Yale U., Zeuthen Workshop–Copenhagen for helpful comments. An earlier version of this paper was circulated under the title 'A Model of Equilibrium Institutions'.

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

Princes who want to make themselves despotic have always begun by uniting all magistracies in their person.

Montesquieu (1748), *De l'esprit des lois*

1 Introduction

Economic efficiency often requires commitment to rules. For example, evidence suggests institutions guaranteeing private property against expropriation are especially important in supporting high levels of economic activity (Acemoglu and Johnson, 2005). A system of private property means a commitment to respect individuals' rights at times when expropriation is physically feasible and in the interests of those in power. The challenge is that this requires restrictions on the untrammelled exercise of power by incumbents — the classic question of 'who will guard the guardians?'

This question might call to mind such notions as 'independent judiciaries', 'the rule of law', 'political representation', and the like. While it would be possible to introduce such devices into a model of political economy by assumption, this would be as a *deus ex machina* that simply overrides the ability of incumbents to exploit their own power. Introducing an exogenous 'higher-level' institutional technology that directly allows for the protection of property rights (perhaps at some cost) would not do justice to the question at hand if the aim is to understand why or why not such commitments arise in equilibrium.

Addressing the question thus requires a framework for studying the struggle for power, which will determine the nature of the prevailing rules. This paper proposes a model where the rules for allocating power and resources are established in the interests of an incumbent group. Rules are contestable in that any group can challenge them through rebellion. There is a simple 'conflict technology' that encompasses all different types of 'rebellions' from 'popular uprisings' to 'coups d'état', with no exogenous restrictions placed on the groups that can fight for power. The defence of the rules rests on the incumbents, who are at an advantage in any fighting, the sense in which they are in power.

Equilibrium rules allocating power and resources are those that maximize the payoff of incumbents, and are stable in that there is no incentive for any group to contest them in favour of alternative equilibrium rules. This means that equilibrium rules are the outcome of a constrained maximization problem taking into account threats of rebellion. It also means a group that rebelled against the former rules would then be similarly constrained by further threats of rebellion. Finally, to discipline the equilibrium concept, equilibrium rules are restricted to depend only on fundamental state variables.

In the model, the rebellion mechanism is the only way to contest the rules, but everyone has access to it. This is important to ensure that results are not due to some individuals having access to other mechanisms to change or resist changes to the various aspects of the rules. In particular, commitment will not arise because of an assumption that some features of an allocation of power and resources cannot ever be changed, nor will commitment fail to arise because some features of an allocation can always be changed without a contest. The model also assumes there are no restrictions on the allocation of resources other than those arising endogenously from the struggle for power. Thus, commitment to property rights will not arise or fail to arise owing to assumptions on the functional forms of feasible tax or transfer schedules.

The paper uses the model outlined above to study the question of whether and how commitment might arise endogenously, and the extent to which the outcome is efficient. There are many examples of time-inconsistency problems where some rules may be in the interests of an incumbent group ex ante, but not ex post. How then can commitment arise if only those with power can put up resistance to changing the rules?

To understand the model, consider first a simple endowment economy with ex-ante identical individuals in which commitment issues will have no bearing on output. Incentives to rebel depend on the payoffs the rebels expect to receive in equilibrium under the new rules as the new incumbents. Since all individuals are ex-ante identical and since the economic environment does not change, the equilibrium rules are the fixed point of the constrained maximization problem of the incumbents subject to the threat of rebellion, where subsequent incumbents would be similarly constrained by equivalent threats of rebellion. As in George Orwell's *Animal Farm*, there is no essential difference between the 'pigs' and the 'men' they replace, but in equilibrium, some individuals will be 'more equal' than others.

In the endowment economy, the model gives rise to a simple theory of distribution. The distribution of resources is uniquely determined and tied to the distribution of power. Those with equal power receive the same payoff, and those with more power receive a higher payoff. The intuition is that in comparing two individuals of equal power, the one with the lower payoff has more to gain from rebellion and is therefore willing to exert greater effort in contesting the allocation; while comparing two individuals with the same payoff, the one with more power poses a greater danger. Since any rebellion will be launched by a subset of the population, the incumbents are concerned with the group of individuals with the greatest incentive to rebel. This means rewarding the powerful to keep them on side, while otherwise equalizing payoffs to avoid concentrating disenchantment. Sharing power thus always entails sharing rents.

There is a basic trade-off in this environment that characterizes the equilibrium rules. On the one hand, the greater the number of individuals sharing power as the incumbent group, the greater their ability to defend the status quo against rebellions, which allows them extract more rents from those outside the group. On the other hand, the rents must then be divided more thinly among more individuals. The equilibrium size of the incumbent group maximizes the payoff of each member by striking a balance between these two effects.

The fully fledged model adds an investment technology to the environment described above, which will give rise to a time-inconsistency problem. Individuals who invest incur an immediate effort cost, while the fruits of their investment are realized only after a lag. During this time, there is the ever-present opportunity for any group of individuals to launch a rebellion. Were a rebellion to occur after investments have been made, the group in power following the rebellion would have incentives to expropriate fully investors' capital because the effort cost of investing would then be sunk.

To provide appropriate incentives for individuals to invest, the rules established prior to investment decisions must offer investors a higher payoff, and importantly, those rules must survive rebellions so that what they prescribe is actually put into practice. In an endowment economy, the incumbents' principal concern is in avoiding a 'popular uprising', a rebellion of outsiders. When offering incentives to investors, the danger of this type of rebellion increases, but it also becomes essential to avoid a 'coup d'état', a rebellion launched by insiders. The higher payoff enjoyed by investors is only in the interests of the incumbents ex ante, so ex post there is a time-inconsistency problem. In other words, the incumbents themselves want to rebel so they can rewrite the rules protecting investors' private property.

It is therefore necessary to reduce the incentive to rebel for those inside and outside the incumbent group simultaneously. This can only be done by expanding the size of the group in power — the problem cannot be solved by any alternative rules established ex ante for the allocation of resources ex post. If higher payoffs were offered to some to reduce their willingness to rebel, resources must be taken away from others, increasing their incentive to rebel. Fundamentally, transfers of resources can only redistribute disgruntlement with the rules. In contrast, enlarging the group in power reduces the attractiveness of all types of rebellions by increasing the number of individuals who will lose power if a rebellion succeeds. This increases the size of the group willing to defend the current rules.

Simply adding the possibility of investment to the model therefore gives rise to an equilibrium with power sharing among a larger group. Sharing power is an *endogenous* commitment mechanism that allows the incumbents to act as a government bound to a set of policies that would otherwise be time inconsistent. The analysis thus highlights the importance of sharing power as a way of guaranteeing commitment to rules, allowing in particular for incentives to invest. This resonates with Montesquieu's doctrine of the separation of powers, now accepted and followed in well-functioning systems of government. Note that power is not shared here with those individuals who are actually investing. The additional individuals in power in no sense represent nor care about those who invest — but they do care about their own rents under the status quo. By this means, a group of self-interested individuals is able to act a government that commits to protection of property rights.

Although it is possible to sustain protection against expropriation by sharing power among a sufficiently large group, in equilibrium, there is too little power sharing and thus too little commitment to protection of property rights. Equilibrium investment is inefficiently low in the sense that total output available for consumption could be increased by having a larger group in power to reduce the proportion of investors' returns that is expropriated. The intuition comes from the inseparability of power and rents, which follows from the threat posed by powerful individuals were conflict to occur. It is not possible in equilibrium for incumbents to share power with more individuals yet not grant them the same payoff as their equally strong peers. This places an endogenous and binding limit on the set of possible transfers, so Pareto-improving deals remain unfulfilled.

The plan of the paper is as follows. There is first a brief review of the literature. Section 2 then presents the model of the power struggle in a simple setting where commitment issues do not arise. The fully fledged model with investment is analysed in section 3. Finally, section 4 draws some conclusions.

1.1 Comparison with the existing literature

This paper shares some features of the literature on coalition formation (see Ray, 2007).[1] As in that literature, the process of establishing rules is non-cooperative, but it is assumed that in the absence of rebellion such rules are followed.[2] Moreover, the modelling of rebellions here is related to blocking in coalitions (Ray, 2007, part III) in the sense that there is no explicit game-form. What distinguishes this paper is the actual modelling of power and conflict. An incumbent coalition proposes a set of rules, but blocking requires (among other things) that an alternative coalition is willing to put in enough fighting effort to displace the current incumbents.

In this sense, this paper is also related to the literature on social conflict and predation, surveyed by Garfinkel and Skaperdas (2007).[3] 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, it would be efficient if individuals could avoid conflict by agreeing and adhering to rules that determine the allocation of resources. This paper supposes such deals are possible. Hence, differently from the literature on conflict, individuals may fight to be part of the group that sets the rules, but not directly over what has been produced. Moreover, they fight in groups, not as isolated individuals.

The idea of sharing power as a commitment device might be reminiscent of the results on the extension of the franchise in Acemoglu and Robinson (2000) and Jack and Lagunoff (2006), but there are two substantial differences here. The first concerns the meaning of sharing power and the mechanism by which it affects economic outcomes. In those papers, extending the franchise means choosing an intrinsically different median voter who will, in future periods, choose policies to which the current elite would like to commit.[4] That type of model assumes the existence of democratic institutions. Here, in contrast, sharing power means simply enlarging the incumbent group (all individuals are ex-ante identical). This does not rely on exogenous differences in preferences, but on changes to the costs of rebellions that in turn affect the rules prevailing in equilibrium. Second, as Acemoglu and Robinson (2000) point out, using England as an example, the extension of the franchise in the second half of the nineteenth century led to a large increase in taxes on the rich

[1] Baron and Ferejohn (1989) analyse bargaining in legislatures using that approach, while Levy (2004) studies political parties as coalitions. Other recent contributions include Acemoglu, Egorov and Sonin (2008), Piccione and Razin (2009), and Acemoglu, Egorov and Sonin (2012).

[2] How would those rules control the allocation of resources 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 individuals coordinate on some pattern of behaviour. One possible interpretation of this approach is similar to the application put forward by Myerson (2009) of Schelling (1960)'s notion of focal points in the organization of society. 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 deviator. 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 of our fellow human beings: those in power choose rules to maximize their own payoffs subject to the threat of rebellions.

[3] See also Grossman and Kim (1995) and Hirshleifer (1995).

[4] See also Bai and Lagunoff (2011).

and a reduction in inequality. By then, the industrial revolution was well underway. The Glorious Revolution, which arguably led to the most decisive break with past government expropriations of property and opened the way for industrialization in England, happened at the end of the seventeenth century and did not coincide with any significant enfranchisement of the poor. Thus, empirically, it is at best unclear that enfranchisement of the poor has a positive effect on protection of property rights.

This paper is also related to work focused on political issues that lead to inefficiencies in protecting property rights, such as Glaeser, Scheinkman and Shleifer (2003), Acemoglu (2008), Guriev and Sonin (2009), and Myerson (2010). Those papers make assumptions on institutions to study how they affect the incentives of judges, oligarchs, dictators, and entrepreneurs, and how that translates into economic outcomes.[5] The objective and the approach of this paper are different: here, assumptions are made on the power struggle to understand whether and how commitment to protect property rights can emerge in equilibrium.

Last, this paper is related to a burgeoning literature on institutions.[6] Recent models of endogenous institutions include Acemoglu and Robinson (2006, 2008), Besley and Persson (2009a,b, 2010), and Greif (2006). Those papers offer many interesting insights, but have not studied how commitment might arise endogenously through power sharing. As discussed in Acemoglu (2003), commitment problems lie at the heart of many political failures (breakdowns of the 'political Coase theorem'), therefore it is important to understand how commitment can arise, which requires a model in which the ability to commit is not assumed but explained.

2 A model of the power struggle

This section presents an analysis of the equilibrium rules allocating power and resources in a simple endowment economy where commitment issues are absent.

2.1 Preferences, technologies, and rules

There is an area containing a measure-one population of ex-ante identical individuals indexed by $\imath \in \Omega$. Individuals receive utility $\mathcal{U}$ from their own consumption C of a homogeneous good and disutility if they exert *fighting effort* F:

$$\mathcal{U} = u(C) - F, \qquad [2.1]$$

[5] Glaeser, Scheinkman and Shleifer (2003) focus on democratic societies and thus assume the existence of a legal system. Acemoglu (2008) compares democratic and oligarchic societies, assuming different institutions in each case. Guriev and Sonin (2009) study the interplay between a dictator and oligarchs, assuming oligarchs can choose between having a weak or a strong dictator (which can be seen as an institution that determines the set of possible actions of the dictator). Myerson (2010) assumes a dictator can choose political liberalization, which is modelled as a probability that the dictator loses power if he expropriates capital.

[6] Many researchers have posited institutions as a significant determinant of economic development and the large differences found in the cross-country distribution of income. For example, see North (1990), North and Weingast (1989), Engerman and Sokoloff (1997), Hall and Jones (1999), and Acemoglu, Johnson and Robinson (2005).

where $u(\cdot)$ is a strictly increasing and weakly concave function.

In this simple economy, individuals who become *workers* receive an exogenous endowment of q units of goods.

There are *rules* stipulating the allocation of power and resources, which will be determined endogenously. These rules specify the set $\mathcal{W}$ of workers, and the set $\mathcal{P}$ of individuals currently in *power*, referred to as the *incumbents*. Each position of power confers an equal advantage on its holder in the event of any conflict, as described below. The term *power sharing* refers to a variable p, defined as the measure of the group $\mathcal{P}$. The incumbent group $\mathcal{P}$ can have any size between 0% and 50% of the population.[7] Those individuals in power cannot simultaneously be workers.[8]

The rules also specify the allocation of goods across all individuals. Let $C_{\mathrm{p}}(\imath)$ and $C_{\mathrm{w}}(\imath)$ denote the individual-specific consumptions of incumbent $\imath \in \mathcal{P}$ and worker $\imath \in \mathcal{W}$ respectively specified by the rules. The rules must satisfy the resource constraint

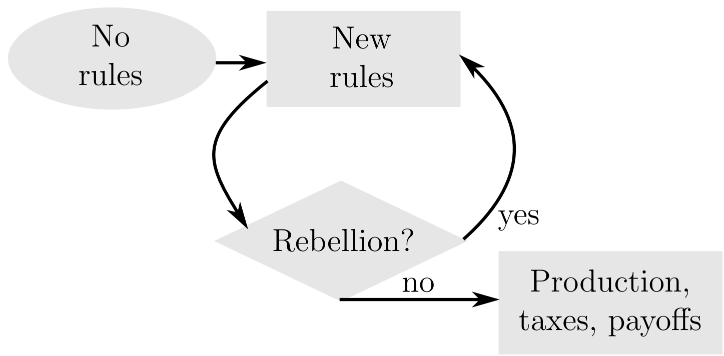
$$\int_{\mathcal{P}} C_{\mathrm{p}}(\imath)\mathrm{d}\imath + \int_{\mathcal{W}} C_{\mathrm{w}}(\imath)\mathrm{d}\imath = \int_{\mathcal{W}} \mathrm{q}\mathrm{d}\imath, \tag{2.2}$$

and each individual's consumption allocation must be non-negative.

Formally, rules are a collection $\mathscr{R} = \{\mathcal{P}, \mathcal{W}, C_{\mathrm{p}}(\imath), C_{\mathrm{w}}(\imath)\}$, where the sets $\mathcal{P}$ and $\mathcal{W}$ satisfy $\mathcal{P} \cup \mathcal{W} = \Omega$, $\mathcal{P} \cap \mathcal{W} = \emptyset$, and $p < 1/2$, and where the consumption allocations $C_{\mathrm{p}}(\imath)$ and $C_{\mathrm{w}}(\imath)$ are consistent with the resource constraint [2.2] and non-negativity constraints.

The sequence of events is depicted in Figure 1. Rules are established. There are then opportunities for *rebellions*, which are described in detail below. If a rebellion occurs, new rules are established, followed by more opportunities for rebellions. When no rebellions occur, workers receive their endowments, the allocation of resources laid down by the prevailing rules is implemented, and payoffs are received.

Figure 1: *Sequence of events*



[7]This is a simple way of modelling decreasing returns in strength with respect to the size of the group in power. The size constraint is generally not binding, but without it, it could be possible that there would be too few individuals outside the incumbent group to launch a rebellion, implying there would be increasing returns to the size of the group in power.

[8]The assumption that those in power do not receive the endowments of workers is not essential for the main results. It does mean that there is an opportunity cost of individuals being in power, so strictly speaking, this is not a pure endowment economy. However, the resulting 'guns versus butter' inefficiency is *not* the focus of this paper.

2.2 Equilibrium rules

Equilibrium rules are those that satisfy three conditions in addition to the basic feasibility constraints set out above. First, they must be in the interests of those in power in the sense of maximizing the average payoff of the incumbents. Second, they must be stable in the sense of avoiding rebellions. Third, they must be Markovian in the sense of depending only on fundamental state variables.

In what follows, let $\mathcal{U}_{\mathrm{p}}$ and $\mathcal{U}_{\mathrm{w}}$ denote the average payoffs of incumbents and workers respectively under a particular set of rules:

$$\mathcal{U}_{\mathrm{p}} = \frac{1}{p}\int_{\mathcal{P}} u(C_{\mathrm{p}}(\imath))\mathrm{d}\imath, \quad \text{and } \mathcal{U}_{\mathrm{w}} = \frac{1}{1-p}\int_{\mathcal{W}} u(C_{\mathrm{w}}(\imath))\mathrm{d}\imath. \qquad [2.3]$$

The first requirement for equilibrium is that the rules maximize the average incumbent payoff $\mathcal{U}_{\mathrm{p}}$ subject to feasibility and the other equilibrium conditions.

The second requirement is that rules are stable, where stability is defined recursively as follows. Rules $\mathscr{R}$ are stable if there are no alternative equilibrium rules $\mathscr{R}' = \{\mathcal{P}', \mathcal{W}', C'_{\mathrm{p}}(\imath), C'_{\mathrm{w}}(\imath)\}$ for which the incumbents under the new rules would be willing to exert sufficient fighting effort to replace the existing rules. Note that the alternative rules $\mathscr{R}'$ must themselves be an equilibrium.

The sufficient fighting effort criterion is now defined. For given alternative rules $\mathscr{R}'$, the fighting effort of those who rebel is sufficient to replace the existing rules if

$$\int_{\mathcal{B}} F(\imath)\mathrm{d}\imath > \int_{\mathcal{A}} \delta\mathrm{d}\imath, \qquad [2.4]$$

where the fighting effort $F(\imath)$ of individual $\imath$, the rebel army $\mathcal{B}$, and the incumbent army $\mathcal{A}$ are given by:

$$F(\imath) = \mathcal{U}'_{\mathrm{p}} - \mathcal{U}(\imath), \quad \mathcal{B} = \{\imath \in \mathcal{P}' \mid \mathcal{U}'_{\mathrm{p}} \geq \mathcal{U}(\imath)\}, \quad \text{and } \mathcal{A} = \{\imath \in \mathcal{P} \mid \imath \notin \mathcal{P}'\ ,\ \mathcal{U}_{\mathrm{p}}(\imath) > \mathcal{U}'_{\mathrm{w}}\}. \qquad [2.5]$$

A *rebellion* is said to occur if [2.4] holds for alternative equilibrium rules $\mathscr{R}'$, in which case the current rules $\mathscr{R}$ are replaced by the new rules $\mathscr{R}'$. The rebel army $\mathcal{B}$ is the set of individuals in the incumbent group $\mathcal{P}'$ under the alternative rules who would gain from this change of the rules. Note that the rebel army can comprise any mixture of current incumbents and workers. The incumbent army $\mathcal{A}$ comprises all individuals in power ($\imath \in \mathcal{P}$) who are not in the rebel army ($\imath \notin \mathcal{P}'$), and who are better off under the current rules than they would be if they lost power under the new rules $\mathscr{R}'$.

A rebellion only occurs if it is known to succeed, and the condition for this is the comparison of fighting strengths in [2.4]. Each army's fighting strength is the integral of the fighting strengths of its members. The fighting strength of individual $\imath \in \mathcal{B}$ in the rebel army is the amount of fighting effort $F(\imath)$ he exerts, which is assumed to equal to the maximum the individual would be willing to put in to change the rules. This is the difference between the individual's anticipated payoff under the post-rebellion rules $\mathcal{U}'_{\mathrm{p}}$ and his payoff under the current rules $\mathcal{U}(\imath)$. The payoff from rebellion is the expected payoff net of fighting effort from being in the incumbent group under the alternative rules, but with the assumption that when there is a non-degenerate distribution of consumption among those in power, individuals do not know in advance where they will be in this distribution.[9] Each individual $\imath \in \mathcal{A}$ in the incumbent army has fighting strength measured by a parameter δ (the power parameter), which is obtained at no utility cost to these individuals.[10]

Although the term *rebellion* has been used to describe the process of changing the rules, the formal definition encompasses 'popular uprisings' (the rebel army comprises only workers), 'coups d'état' (the rebel army is a subset of the incumbent group), 'suspensions of the constitution' (the rebel army includes all incumbents), and 'revolutions' that receive the backing of some insiders from the current regime (the rebel army includes a mixture of workers and incumbents).

The third requirement for rules to be an equilibrium is that they are Markovian. Formally, the sizes of the groups of workers and incumbents, and the distributions of consumption among these individuals can depend only on fundamental state variables and not on the general history of earlier play.

Since individuals are ex-ante identical, there is an essential indeterminacy in the identities of the incumbents and in the assignment of particular consumption levels to incumbents and workers in the case of non-degenerate distributions. If a certain set of rules $\mathscr{R}$ is an equilibrium, otherwise identical rules with any permutation of identities will also be an equilibrium. Therefore, in the constrained maximization problem that the equilibrium rules must satisfy, the variables that are determined are the extent of power sharing p, and distributions of consumption for incumbents and workers, denoted by $C_{\mathrm{p}}(\cdot)$ and $C_{\mathrm{w}}(\cdot)$ respectively, which make no reference to identities.

The conditions for rules to be an equilibrium are collected in the following definition.

Definition 1 *Rules $\mathscr{R}^* = \{\mathcal{P}^*, \mathcal{W}^*, C^*_{\mathrm{p}}(\imath), C^*_{\mathrm{w}}(\imath)\}$ are an equilibrium if the following conditions are satisfied:*

1. *[**Optimal**] Power sharing p^* and distributions $C^*_{\mathrm{p}}(\cdot)$ and $C^*_{\mathrm{w}}(\cdot)$ of consumption for incumbents and workers maximize the average utility $\mathcal{U}_{\mathrm{p}}$ of incumbents.*

2. *[**Stable**] There are no rules $\mathscr{R}' = \{\mathcal{P}', \mathcal{W}', C'_{\mathrm{p}}(\imath), C'_{\mathrm{w}}(\imath)\}$ such that $\mathscr{R}'$ is itself an equilibrium and the successful rebellion condition [2.4] holds.*

3. *[**Markovian**] Power sharing p^* and distributions $C^*_{\mathrm{p}}(\cdot)$ and $C^*_{\mathrm{w}}(\cdot)$ of consumption for incumbents and workers depend only on fundamental state variables.* □

2.3 Characterizing the equilibrium rules

In the simple endowment-economy model, there are no fundamental state variables in the problem of finding the optimal $\{p, C_{\mathrm{p}}(\cdot), C_{\mathrm{w}}(\cdot)\}$. The equilibrium values of p^*, $C^*_{\mathrm{p}}(\cdot)$ and $C^*_{\mathrm{w}}(\cdot)$ are then found in two steps. First, the value of p and the distributions $C_{\mathrm{p}}(\cdot)$ and $C_{\mathrm{w}}(\cdot)$ are chosen to maximize the average payoff of those in power subject to the stability condition, where p', $C'_{\mathrm{p}}(\cdot)$ and $C'_{\mathrm{w}}(\cdot)$ are

[9] In equilibrium, there will be a degenerate consumption distribution within the incumbent group, but that assumption simplifies the analysis.

[10] Section 2.4 discusses the assumptions on the power struggle underlying this rebellion mechanism.

taken as given. Although the stability condition is stated in terms of sets $\mathcal{P}$ and $\mathcal{P}'$ and allocations $C_{\mathrm{p}}(\imath)$ and $C_{\mathrm{w}}(\imath)$, the implied restrictions on power sharing p and the distributions $C_{\mathrm{p}}(\cdot)$ and $C_{\mathrm{w}}(\cdot)$ do not depend on the identities of the individuals in the sets $\mathcal{P}$ and $\mathcal{P}'$. The stability condition is basically a set of *no-rebellion constraints* for each possible composition of the rebel army, taking into account differences in individuals' consumption allocations and power. Second, the equilibrium conditions of identical power sharing $p' = p^*$ before and after rebellions and identical distributions of consumption for both groups $C'_{\mathrm{p}}(\cdot) = C^*_{\mathrm{p}}(\cdot)$ and $C'_{\mathrm{w}}(\cdot) = C^*_{\mathrm{w}}(\cdot)$ are imposed.

The following proposition derives some necessary features of any equilibrium rules in this environment.

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

(i) *Equalization of workers' payoffs:* $\mathcal{U}^*_{\mathrm{w}}(\imath) = \mathcal{U}^*_{\mathrm{w}}$ *for all* $\imath \in \mathcal{W}$ *(with measure one).*

(ii) *Sharing power implies sharing rents:* $\mathcal{U}^*_{\mathrm{p}}(\imath) = \mathcal{U}^*_{\mathrm{p}}$ *for all* $\imath \in \mathcal{P}$ *(with measure one).*

(iii) *Power determines rents:* $\mathcal{U}^*_{\mathrm{p}} - \mathcal{U}^*_{\mathrm{w}} = \delta$.

(iv) *The stability condition can be reduced to a single 'no-rebellion' constraint for a rebel army including only workers:*

$$\mathcal{U}'_{\mathrm{p}} - \mathcal{U}_{\mathrm{w}} \leq \delta \frac{p}{p'}, \qquad [2.6]$$

with all other 'no-rebellion' constraints being redundant.

(v) *An equilibrium always exists and is unique (up to permutations of identities). Equilibrium power sharing* p^* *satisfies* $0 < p^* \leq 2 - \varphi$, *where* $\varphi \equiv (1 + \sqrt{5})/2 \approx 1.618$.[11]

Proof See appendix A.1. ■

The first two parts of the proposition demonstrate that incumbents have a strong incentive to avoid inequality except where it is matched by differences in power. These results hold even when the utility function is linear in consumption, and so do not depend on strict concavity.

The intuition for the payoff-equalization results is that the most dangerous composition of the rebel army is the one including those individuals with the greatest incentive to fight. A rebel army will always be a subset of the whole population. As a consequence, if there were payoff inequality among workers, the most dangerous rebel army would not include those workers who receive a relatively high payoff. The incumbents could then reduce the effectiveness of this rebel army by redistributing from relatively well-off workers to the worse off. This slackens the set of no-rebellion constraints, allowing the incumbents to achieve a higher payoff. These gains are exploited to the maximum possible extent when all workers' payoffs are equalized.[12] Similarly, inequality in

[11] The term $\varphi \equiv (1 + \sqrt{5})/2$ is known as the *Golden ratio* or *mean of Phidias*. The constraint $p < 1/2$ is always slack in equilibrium.

[12] 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.

incumbent payoffs is undesirable because incumbents receiving a relatively low payoff can defect and join a rebel army. Equalizing incumbent payoffs by redistributing consumption does not directly lower their average utility when the utility function is weakly concave, while it has the advantage of weakening the most dangerous rebel army. Since defections from the group in power weaken the incumbent army, there is no version of this argument that calls for equalization of payoffs *between* workers and incumbents.

Making use of the payoff-equalization results and the resource constraint, the utilities of workers and incumbents are:

$$\mathcal{U}_{\mathrm{w}} = u(C_{\mathrm{w}}), \quad \text{and } \mathcal{U}_{\mathrm{p}} = u(C_{\mathrm{p}}) = u\left(\frac{(1-p)(q-C_{\mathrm{w}})}{p}\right). \qquad [2.7]$$

As a consequence of the payoff-equalization results, all that matters for the composition of a rebel army are the fractions σ_{w} and σ_{p} of its total numbers drawn from workers and from the group in power. The equilibrium rules must then be a solution of the problem

$$\max_{p,C_{\mathrm{w}}} \mathcal{U}_{\mathrm{p}} \text{ s.t. } \sigma_{\mathrm{w}} \max\{\mathcal{U}'_{\mathrm{p}} - \mathcal{U}_{\mathrm{w}}, 0\} + \sigma_{\mathrm{p}}(\mathcal{U}'_{\mathrm{p}} - \mathcal{U}_{\mathrm{p}} + \delta)\mathbb{1}[\mathcal{U}'_{\mathrm{p}} \geq \mathcal{U}_{\mathrm{p}}] \leq \delta\frac{p}{p'}, \qquad [2.8]$$

for all non-negative proportions σ_{w} and σ_{p} that are feasible given the size of the rebel army p' and the sizes of the groups of workers and incumbents under the current rules, that is, $\sigma_{\mathrm{w}} \leq (1-p)/p'$, $\sigma_{\mathrm{p}} \leq p/p'$, and $\sigma_{\mathrm{w}} + \sigma_{\mathrm{p}} = 1$. The general form of the no-rebellion constraints stated in [2.8] is derived from the participation constraints on membership of the rebel and incumbent armies described in equations [2.3]–[2.4],[13] with $\mathbb{1}[\cdot]$ denoting the indicator function.

The fourth claim in Proposition 1 states that the equilibrium rules can be characterized by a single 'no-rebellion' constraint [2.6], which is equivalent to setting $\sigma_{\mathrm{w}} = 1$ and $\sigma_{\mathrm{p}} = 0$ in the general constraints of [2.8] (and noting that $\mathcal{U}'_{\mathrm{p}}$ will exceed $\mathcal{U}_{\mathrm{w}}$ in equilibrium). Satisfaction of [2.6] is clearly necessary, but the proposition shows that this single constraint is also *sufficient* to characterize the equilibrium rules.[14] The constrained maximization problem thus reduces to

$$\max_{p,C_{\mathrm{w}}} \mathcal{U}_{\mathrm{p}} \text{ s.t. } \mathcal{U}'_{\mathrm{p}} - \mathcal{U}_{\mathrm{w}} \leq \delta\frac{p}{p'}, \qquad [2.9]$$

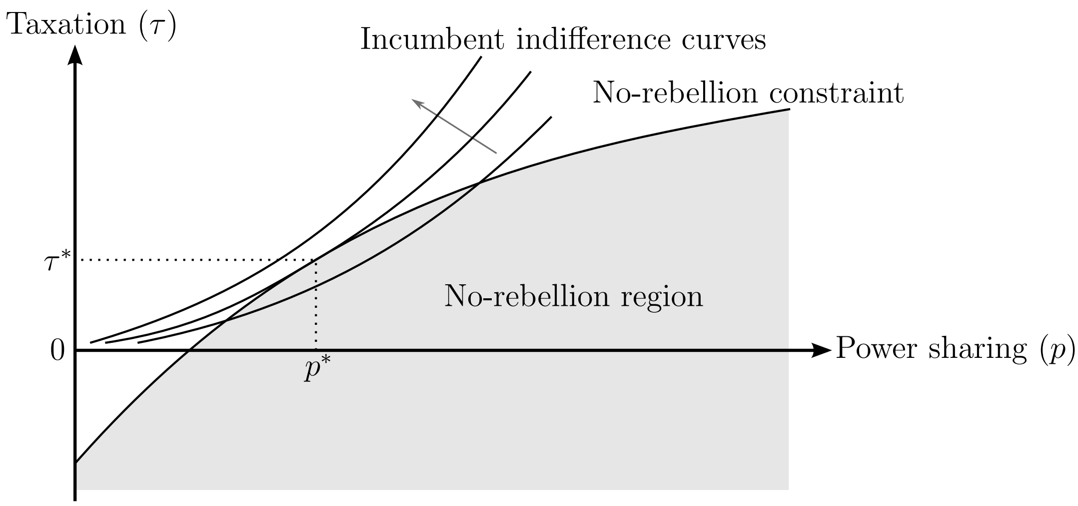
where $\mathcal{U}_{\mathrm{w}}$ and $\mathcal{U}_{\mathrm{p}}$ are as in equation [2.7], with p' and $\mathcal{U}'_{\mathrm{p}}$ taken as given. In equilibrium, these are equal to the corresponding values p^* and $\mathcal{U}^*_{\mathrm{p}}$ that solve the maximization problem.

The equilibrium allocation of resources can be implemented by having each worker $\imath \in \mathcal{W}$ face a tax (or transfer if negative) $\tau = \mathrm{q} - C_{\mathrm{w}}$. The payoff of each incumbent is increasing in the tax τ, and decreasing in power sharing p because the total tax revenue must be distributed more widely (and because there are fewer workers to tax). The indifference curves of the incumbents over τ and p and the single binding no-rebellion constraint are plotted in Figure 2.

[13] The requirement $\mathcal{U}_{\mathrm{p}} > \mathcal{U}'_{\mathrm{w}}$ that incumbents who do not join a rebel army are willing to defend the current rules is always satisfied in equilibrium.

[14] This finding is specific to the simple endowment-economy model of this section. In the fully fledged model with investment, other no-rebellion constraints become binding. Which compositions of the rebel army are associated with

Figure 2: *Trade-off between power sharing and taxation*



An increase in τ reduces the payoffs of workers, making them more willing to fight in a rebellion, while an increase in the size of the group in power boosts the fighting strength of the incumbent army, making rebellion less attractive to workers. The incumbents thus have two margins to avoid rebellions. They can reduce taxes τ (the 'carrot'), or increase their size p (the 'stick'). This corresponds to the upward-sloping 'no-rebellion' constraint depicted in Figure 2.[15]

After taking into account the binding no-rebellion constraint, the key decision is how widely to share power. Incumbents face a fundamental trade-off in determining their optimal size: a larger size will make rebellions more costly and allow higher taxes to be extracted from workers, but will also spread the revenue from these taxes more thinly among a larger number of individuals (and also reduce the tax base). Proposition 1 shows that it is not possible in equilibrium to include extra individuals in the group in power without offering these individuals the same high payoff received by other members. Thus, sharing power entails sharing rents. The incumbents then share power with an extra individual if and only if this allows them to increase their average payoff. The allocation of power therefore reflects the interests of the incumbents, rather than the interests of society. In equilibrium, the utility value of the rents received by those in power depends only on the exogenous parameter δ.

2.4 Discussion

The model described above is designed to capture in a simple manner the 'power struggle' for control of the allocation of resources. The rules for allocating power and resources can be overthrown by rebellions and replaced by new ones. The success of a rebellion is settled by a basic 'conflict technology' that avoids going into the punches and sword thrusts of battle. Everyone has access on the same terms to opportunities for rebellion, irrespective of their current status. Several assumptions

binding no-rebellion constraints in equilibrium is endogenous and will depend on the situation being analysed.

[15]When utility is linear in consumption, the no-rebellion constraint is a straight line. The diagram shows the general case where utility is strictly concave in consumption, resulting in the constraint being a concave function.

are made for simplicity. The combatants' strengths are linear in the strengths of the individuals that make up the rebel and incumbent armies. Members of the incumbent army have a predetermined fighting strength (the power parameter δ), so this is inelastic with respect to current fighting effort. Given fighting strengths, there is no uncertainty about which side will emerge victorious.

In modelling conflict, it is necessary to introduce some asymmetry between the rebel and incumbent armies for the notion of being 'in power' to be meaningful. The parameter δ represents the advantage incumbents derive from their entrenched position. It is fighting strength that is obtained at no *current* effort cost (though past effort may have been expended in becoming an incumbent), while the rebels can only obtain fighting strength from current effort. The inelasticity of incumbent-army members' fighting strength with respect to current effort can be seen as an inessential simplification.[16] When the current rules are established, there are several margins that can be adjusted to avoid rebellions, such as varying the number of individuals in power, or increasing the consumption allocated to those who would otherwise join a rebel army. The ability to adjust each incumbent's power δ at some cost in addition to these does not fundamentally change the problem. The rebel army, however, lacks these alternative margins, so it is essential that its fighting strength is responsive to the effort put in by its members.

One interpretation of the parameter δ is that the individuals currently in power possess some defensive fortifications, such as a castle, which place them at an immediate fighting advantage over any rebels who must breach these from outside. A broader interpretation is that δ represents the more severe coordination problems faced by rebels from outside the incumbent group. Authority depends on a chain of command, where individuals follow instructions in expectation of punishment from others if they disobey. The rebels confront the challenge of persuading enough individuals that they should fear punishment for disobeying them rather than the incumbents people are accustomed to.[17]

The assumptions of the model allow for coordination among individuals in launching rebellions, but subject to some restrictions. These restrictions are intended as a simple representation of the plausible constraints that ought to be placed on the set of possible deals or 'contracts' among the rebels. The fundamental contracting problem is the issue of enforceability in a world with no exogenous commitment technologies. The rebellion mechanism is intended to be as flexible as possible in allowing for 'deals' subject to the limits of enforceability.

The rebellion 'contract' implicit in the model requires a prescribed amount of fighting effort from each rebel in return for a place in the group in power under the new rules. There is a restriction that rebel armies are open only to those who expect a place in the group subsequently in power. This rebellion contract raises two questions. Why are other forms of contract ruled out? And what suggests a contract of this form is not susceptible to enforceability concerns?

[16] The inelasticity assumption can also be viewed as a requirement of internal consistency given the deterministic nature of the conflict technology. If the incumbents knew they would be defeated, it is difficult to rationalize their exerting any current fighting effort.

[17] Under this interpretation, the successful rebellion condition [2.4] can be seen as the rebels' effort requirement to demonstrate that they have the strength and the organization to overcome these problems. Once the incumbents see this tipping point is reached, they surrender and no actual fighting takes place.

Consider a group of individuals with incentives to come together and agree to fight, enabling new rules to be established that offer each one of them a better payoff than what they currently receive. The maximum fighting effort each would be willing to agree to is equal to each's expected utility gains. There are two facets of the enforceability problem for this 'deal'. First, after the success of the rebellion, will there be incentives to establish the particular rules that were agreed beforehand? Second, will individual rebels honour their agreed levels of fighting effort?

A rebellion contract prescribing exactly what rules are to be established faces formidable enforceability problems. Once the status quo is overthrown, there is no higher authority that can compel the group now in power to act against its interests ex post. This type of contract is therefore ruled out. The new rules for allocating power and resources must maximize the payoffs of those who are now the incumbents[18] starting afresh from the world as they find it, unconstrained by history except for fundamental state variables.[19] Any bygones will be bygones. In particular, this precludes the trigger strategies of repeated games as commitment devices.[20]

As Proposition 1 shows, the incumbents have a strict preference to avoid payoff inequality except where it is matched by differences in power. It follows that the group now in power would have incentives not to honour contracts that specified either transfers to those who contributed fighting effort during the rebellion or fines for those who did not, hence such contracts cannot be written ex ante.

Now consider the second facet of the enforceability problem. Taking as given the payoff improvement an individual expects if a rebellion succeeds (subject to the restrictions on what can be agreed in advance regarding the new rules), will it be possible to enforce the agreed amount of fighting effort from an individual who is a party to a rebellion contract? The basic problem is that each atomistic individual (correctly) does not perceive himself as pivotal in determining whether the rebellion succeeds. Thus, left to his own devices, he would have an incentive to shirk and free-ride on others' fighting effort. To a large extent, rebel armies may be able to control individual members through internal discipline, but a non-negligible enforceability problem remains when some necessary fighting is done at an individual's discretion.

To ensure that all agreed fighting effort is actually exerted, there needs to be a credible punishment that can be imposed on shirkers after the fact. However, according to Proposition 1, only differences in payoffs that reflect differences in power are in the interests of the incumbents ex post. This suggests that the offer of a position of power conditional on the requisite amount of fighting effort can provide a credible incentive not to shirk.[21] While the rebellion contract cannot determine

[18] In the model, rules being in the interests of the incumbent group is interpreted to mean maximizing their average payoff. Moving away from this simplifying assumption would require modelling the hierarchy and power relations within the incumbent group. See Myerson (2008) for a model which addresses that question.

[19] For simplicity, the occurrence of conflict does not itself affect any fundamental state variables. This implicitly assumes members of the rebel army can be demobilized costlessly once the fighting is over if, off the equilibrium path, there were more rebels than places in the new incumbent group. Adding a cost of demobilization would make the size of the previous rebellion a state variable at the stage new rules are established, which would add a significant complication to the model without obviously delivering any new insights.

[20] Models that allow trigger strategies face the problem of multiple equilibria because there is always a range of possible punishments consistent with equilibrium.

[21] Suppose that a fraction ξ of an individual's agreed fighting effort (associated with some necessary tasks) cannot

the total number of positions of power under the new rules, it can control the identities of those who will receive these positions.[22] But for those who anticipate becoming workers, even if they gain from the success of the rebellion, there is no worse position they can be credibly assigned if they fail to put in the agreed level of fighting effort. There is nothing to prevent such individuals from shirking,[23] which leads to the restriction that all rebels must expect a position of power.

2.5 Example I: utility linear in consumption

There are three exogenous parameters in the model: the power parameter δ of an incumbent, the endowment q of a worker, and the utility function $u(\cdot)$ in consumption. This example illustrates the workings of the model with a linear utility function $u(C) = C$. The maximization problem [2.9] becomes

$$\max_{p, C_{\mathrm{w}}} \frac{(1-p)(q-C_{\mathrm{w}})}{p} \quad \text{s.t.} \quad C'_{\mathrm{p}} - C_{\mathrm{w}} \leq \delta \frac{p}{p'}, \tag{2.10}$$

after substituting the expressions for $\mathcal{U}_{\mathrm{p}}$ and $\mathcal{U}_{\mathrm{w}}$ from [2.7]. The single no-rebellion constraint is binding, and can be used to solve explicitly for the consumption of a worker $C_{\mathrm{w}} = C'_{\mathrm{p}} - \delta p/p'$. Substituting that into the objective function yields the consumption of an incumbent:

$$C_{\mathrm{p}} = \frac{1-p}{p}\left(\mathrm{q} - C'_{\mathrm{p}} + \delta \frac{p}{p'}\right). \tag{2.11a}$$

The problem is now an unconstrained choice of power sharing p to maximize each incumbent's consumption, with p' and C'_{p} taken as given. The first-order condition is

$$\frac{C^*_{\mathrm{p}}}{1-p^*} = (1-p^*)\frac{\delta}{p'}. \tag{2.11b}$$

be directly enforced at the time through the rebel army's own discipline mechanisms, but that the full amount of fighting effort F must be exerted otherwise the individual does not obtain fighting strength F. Suppose also that each individual's total fighting effort is verifiable after the rebellion. After agreeing to the rebellion contract, the individual is directly compelled to exert effort $(1-\xi)F$. If he exerts the remaining effort ξF then he subsequently receives his position of power with continuation payoff $\mathcal{U}'_{\mathrm{p}}$. If he shirks and exerts no further effort, he is demoted to worker status and receives payoff $\mathcal{U}'_{\mathrm{w}}$. Therefore, for individual $\imath$ to join the rebel army and contribute to the fighting, incentive compatibility requires $\mathcal{U}'_{\mathrm{p}} - \mathcal{U}'_{\mathrm{w}} \geq \xi F(\imath)$, while the maximum-effort participation constraint is $\mathcal{U}'_{\mathrm{p}} - \mathcal{U}(\imath) \geq F(\imath)$. If ξ is positive but sufficiently small then the incentive compatibility constraint is satisfied but not binding, while the participation constraint binds, as was implicitly assumed in the description of the rebellion mechanism.

[22] As the identities of the incumbent group have no effect on the maximum attainable payoffs of those in power, carrying out the punishment would not affect the others in the group. Intuitively, since all individuals are ex ante identical, individuals in power do not care about the identities of those with whom they share power, only the total number of such people.

[23] The incentive compatibility constraint discussed in footnote 21 would be violated for these individuals with any positive discretionary effort ξ, no matter how small. One alternative approach that could incentivize more individuals not to shirk offers a lottery in return for an agreed amount of fighting effort, where the prize is a position of power. While this mechanism could induce fighting effort from more individuals, the amount of effort each individual would agree to is lower because the lottery is less valuable than a position of power with certainty.

The equilibrium conditions ($p^* = p'$, $C_\mathrm{p}^* = C_\mathrm{p}'$) are now imposed in [2.11a], which leads to $C_\mathrm{p}^* = (\mathrm{q} + \delta)(1 - p^*)$. Combining this with equation [2.11b] (and using $p^* = p'$ again) yields the equilibrium:[24]

$$p^* = \frac{\delta}{\mathrm{q} + 2\delta}, \quad C_\mathrm{p}^* = \frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta}, \quad \text{and} \;\; C_\mathrm{w}^* = \frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta} - \delta. \qquad [2.12]$$

Notice in this case that power sharing is a function of the ratio δ/q.[25]

2.6 Example II: public goods

In the model so far, there is no scope for incumbents to do what governments are customarily supposed to do, such as provide public goods. This example extends the model by introducing a technology that allows for production of public goods. The rules for allocating power and resources now also specify spending on public goods. It is then natural to ask whether resources will be efficiently allocated to public-good production.

The new technology converts units of output into public goods. If g units of goods per person are converted then everyone receives an extra $\Gamma(g)$ units of the consumption good. The function $\Gamma(\cdot)$ is strictly increasing, strictly concave, and satisfies the usual Inada conditions. The definition of rules $\mathscr{R}$ from section 2.2 is augmented to specify public-good provision g, hence $\mathscr{R} = \{\mathcal{P}, \mathcal{W}, C_\mathrm{p}(\imath), C_\mathrm{w}(\imath), g\}$. All individuals observe the choice of g and take it into account — along with all other aspects of the rules — when deciding whether to participate in a rebellion, and if so, the amount of fighting effort to exert. There is no other change to the environment.

The utility of an individual is now

$$\mathcal{U} = u(C) - F, \quad \text{with} \;\; C = c + \Gamma(g), \qquad [2.13]$$

where C denotes the individual's overall consumption, comprising private consumption c and the consumption $\Gamma(g)$ each person obtains from the public good. The resource constraint is now

$$pC_\mathrm{p} + (1-p)C_\mathrm{w} = (1-p)\mathrm{q} - g + \Gamma(g). \qquad [2.14]$$

A benevolent social planner would choose the first-best level of public-good provision $g = \hat{g}$, determined by $\Gamma_g(\hat{g}) = 1$, which maximizes the total amount of goods available for consumption.

[24]The parameter restriction $\delta/\mathrm{q} \leq \varphi$ is assumed, where φ is the Golden ratio (see footnote 11). When the utility function is linear, this restriction is necessary and sufficient for an equilibrium in which the non-negativity constraint on workers' consumption is not binding.

[25]The power parameter δ affects the equilibrium in three ways. First, an increase in δ makes the incumbents stronger because the rebels have to exert greater fighting effort to defeat the incumbent army. This 'income effect' leads to a reduction in C_w and a decrease in p. Second, the payoff that the rebels would receive once in power increases because their position will also be stronger once they have supplanted the current incumbents, making rebellion more attractive. This offsetting 'income effect' increases C_w and increases p. Third, an increase in δ raises the effectiveness of the marginal fighter in the incumbent army, leading to a 'substitution effect' whereby the incumbents increase their size in order to extract higher rents. As long as the non-negativity constraint on workers' consumption is not binding, the third effect dominates and power sharing is increasing in δ.

In determining the equilibrium rules, the payoff-equalization insights of Proposition 1 continue to apply to this new environment, hence it is possible without loss of generality to focus on rules specifying power sharing p, the consumption C_{w} of all workers, and the necessarily common public-good provision g. The resource constraint [2.14] can be used to find the consumption level of each incumbent under a particular set of rules:

$$C_{\mathrm{p}} = \frac{(1-p)(q-C_{\mathrm{w}})}{p} + \frac{\Gamma(g)-g}{p}. \qquad [2.15]$$

The argument of Proposition 1 that the equilibrium rules can be characterized by imposing only the no-rebellion constraint for a rebel army composed entirely of workers also carries over to this new environment. Thus, the equilibrium rules are the solution of

$$\max_{p,C_{\mathrm{w}},g} u\left(\frac{(1-p)(q-C_{\mathrm{w}})}{p} + \frac{\Gamma(g)-g}{p}\right) \text{ s.t. } \mathcal{U}'_{\mathrm{p}} - u\left(C_{\mathrm{w}}\right) \leq \delta\frac{p}{p'}, \qquad [2.16]$$

with p' and $\mathcal{U}'_{\mathrm{p}}$ taken as given, but with $p'=p^*$, $C'_{\mathrm{w}}=C^*_{\mathrm{w}}$ and $g'=g^*$ in equilibrium. The first-order condition for public-good provision g is:

$$\Gamma_g(g^*) = 1. \qquad [2.17]$$

This is identical to the condition for the provision $\hat{g}$ chosen by a benevolent social planner, so $g^*=\hat{g}$. The equilibrium rules are therefore economically efficient in respect of public-good production.[26]

To understand this result, observe that the no-rebellion constraint implies the incumbents cannot disregard the interests of the workers, even though they do not care about them directly. Provision of the public good slackens the no-rebellion constraint, while the resources appropriated to finance it tighten the constraint. By optimally trading off the benefits of the public good against the cost of production, the incumbents effectively maximize the size of the pie, making use of transfers to ensure everyone is indifferent between rebelling or not. This efficiency result can be seen as a 'political' analogue of the Coase theorem,[27] where the contestability of rules through rebellions plays the role of legal property rights. The ability of rules to determine the allocation of resources is crucial to this finding, but more importantly, the constraints imposed by the power struggle do not interfere with the transfers needed here.

The analysis shows that although the incumbents are extracting rents from workers, this does not preclude them from acting as if they were benevolent in other contexts. Hence, the overall welfare of workers might be larger or smaller compared to a world in which no-one can compel others to act against their will. This reflects the ambivalent effects on ordinary people of having a ruling elite.[28]

[26]The distribution of total output between workers and incumbents depends on the other parameters of the model, including the utility function $u(\cdot)$. In equilibrium, all individuals will receive a higher overall payoff as a result of the public-good technology being available, though in general, the benefits will not be shared equally.

[27]See Acemoglu (2003) for a discussion of this analogy.

[28]This trade-off is mentioned in the Bible (1 Samuel 8:10–20). The people want a king to provide public goods, despite being warned by the prophet Samuel that the king would use his power in his own interests. Many centuries later, in far too many cases, the warnings of Samuel remain as relevant as ever.

The result found with this example is far from surprising and can be obtained in several other settings, as discussed by Persson and Tabellini (2000) in the context of voting and elections. Here the result provides a benchmark case where the equilibrium rules are economically efficient.

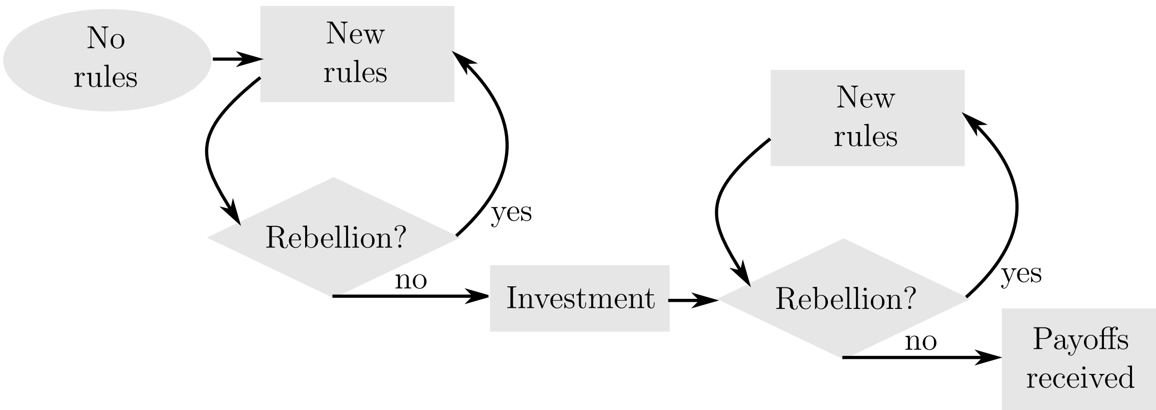
3 Investment

This section adds the possibility of investment to the analysis of equilibrium rules. Individuals can now exert effort to obtain a greater quantity of goods, but there is a time lag between the effort being made and the fruits of the investment being received. During this span of time, there are opportunities for rebellion against the prevailing rules. The model is otherwise identical to that of section 2. In particular, there are no changes to the mechanism through which rules are established and changed. However, the occurrence of investment changes incentives for rebellion, and thus affects the constraints the equilibrium rules must satisfy. The following analysis considers how the equilibrium rules will be able to provide credible incentives for individuals to invest, and to what extent it will be done — in particular, whether these rules will achieve economic efficiency.

3.1 Environment

The sequence of events is depicted in Figure 3. Before any investment decisions are made, rules are first established through a process identical to that described in section 2 (compare Figure 1). The rules specify a consumption allocation, which can now be contingent on individuals' investment decisions, as well as determine the size of the group in power. Once rules that do not immediately trigger rebellion are established, there are opportunities to invest. After investment decisions are made, there is another round of opportunities for rebellion, with new rules established if a rebellion occurs.

Figure 3: *Sequence of events with investment*



Individuals who are in power (incumbents, denoted by $\imath \in \mathcal{P}$) have fighting strength δ in defence, as in the model of section 2. Individuals not belonging to the incumbent group (denoted by $\imath \in \mathcal{N}$) at the post-investment stage receive an endowment of q units of goods.

There are μ *investment opportunities*. An investment opportunity is the option to produce κ units of capital in the future in return for incurring a present effort cost θ (in utility units), which is sunk by the time the capital is produced. *Capital* here simply means more units of the consumption good. All investment opportunities yield the same amount of capital κ, but each features an effort cost that is an independent draw from the distribution

$$\theta \sim \text{Uniform}\,[\psi, \kappa]\,, \tag{3.1}$$

where $0 < \psi < \kappa$, with ψ being the minimum effort cost.[29] An individual's receipt of an investment opportunity, the required effort cost θ, and whether the opportunity is taken, are private information at the time of the investment decision, while the individual's production of capital becomes common knowledge after the investment stage.[30] It is further assumed that investment opportunities are received only by those individuals outside the incumbent group,[31] and that investment opportunities are randomly assigned at the investment stage with no-one having prior knowledge of whether he will receive one, nor the required effort cost θ if so.[32]

An individual's utility function is

$$\mathcal{U} = u(C) - \theta I - F, \tag{3.2}$$

where $I \in \{0, 1\}$ is an indicator variable for whether an investment opportunity is received and taken, and F denotes any fighting effort, as in the model of section 2.

The following parameter restrictions are imposed:

$$\frac{\delta}{q} \leq \varphi \equiv \frac{1+\sqrt{5}}{2}, \quad \mu < \frac{q}{2(q+2\delta)}, \quad \kappa < \delta, \quad \text{and} \;\; u(C) = C. \tag{3.3}$$

The first restriction is the bound from section 2.5 needed to ensure non-negativity constraints are always slack in equilibrium. The second restriction states that the measure μ of individuals who receive an investment opportunity is not too large, which ensure there is never a shortage of workers to fill a rebel army after investment has occurred.[33] The third restriction places a physical limit

[29] The uniform distribution is chosen for simplicity. The choice of distribution does not affect the qualitative results.

[30] Since investment opportunities are private information when taken, it is not feasible to have rules specifying a 'command economy' where individuals perform investments by decree.

[31] Allowing those in power to invest adds extra complications to the model. It might be thought important to have investors inside the incumbent group to provide appropriate incentives. As will be seen, this is not the case, and moreover, the advantage of having investors in power most likely applies to the case where they are brought into the incumbent group *conditional* on taking an investment opportunity. From the incumbents' perspective, the incentive to do this disappears once investments have come to fruition. But given the information restrictions, which represent the not implausible difficulties of identifying talented investors in advance, bringing in individuals conditional on investment at an earlier stage is not feasible.

[32] This modelling device places individuals behind a 'veil of ignorance' about their talents as investors when the pre-investment stage rules are determined. Doing this avoids having to track whether talented investors are disproportionately inside or outside the group in power, which would add a (relevant) state variable to the problem of determining the pre-investment stage rules, significantly complicating the analysis. However, it will turn out that the no-rebellion constraints are slack at the pre-investment stage, so this assumption need not significantly affect the results.

[33] If capitalists were to become the predominant group then the nature of the binding constraints might change and

on the economy's maximum capital stock. Finally, individuals' utility functions are assumed to be linear in consumption for analytical tractability. This allows for a simple closed-form solution, but it is expected that similar results would be found for the general class of concave utility functions.

In this environment, rules are formally defined as $\mathscr{R} = \{\mathcal{P}, \mathcal{N}, C_{\mathrm{p}}(\imath), C_{\mathrm{w}}(\imath), C_{\mathrm{k}}(\imath)\}$, where $\mathcal{P}$ is the set of incumbents and $\mathcal{N}$ is the set of all other individuals. The consumption allocation of the incumbents is denoted by $C_{\mathrm{p}}(\imath)$. For an individual $\imath \in \mathcal{N}$ outside the incumbent group, the consumption allocation is potentially contingent on whether the individual has produced capital. The term *worker* now refers to an individual who did not produce any capital, in which case the consumption allocation is $C_{\mathrm{w}}(\imath)$ for individual $\imath$. Those who did produce capital, the *capitalists*, receive consumption allocation $C_{\mathrm{k}}(\imath)$.

The notion of equilibrium rules is unchanged from that introduced in Definition 1. Equilibrium rules must be optimal from the perspective of the incumbents, and they must be stable and Markovian. Stability of rules established at the pre-investment stage now requires the absence of rebellions both before and after the investment stage. At the pre-investment stage, because those outside the incumbent group do not know their type yet, their payoffs are to be interpreted as expected payoffs. At the post-investment stage, continuation payoffs are independent of any earlier effort (this is the sense in which effort costs are sunk). Finally, since the capital stock is a fundamental state variable at the post-investment stage, the Markovian condition allows rules established at that stage to depend on capital.

Since equilibrium rules would remain so after a permutation of identities, the relevant variables characterizing the rules are power sharing p, and the distributions $C_{\mathrm{p}}(\cdot)$, $C_{\mathrm{w}}(\cdot)$, and $C_{\mathrm{k}}(\cdot)$ of consumption for incumbents, workers, and capitalists respectively.

3.2 Characterizing the equilibrium rules

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

3.2.1 Post-investment stage rules after a rebellion

Suppose that a rebellion occurs at some point after investment decisions have been made, which is necessarily off the equilibrium path given the stability condition for equilibrium rules at the pre-investment stage. Let K denote the predetermined stock of capital that has been produced. The effort costs θ of investing are now sunk, so the continuation utility function $\mathcal{U} = C - F$ is the same for both an expropriated capitalist and an individual who never possessed any capital to begin with. The resource constraint at the post-investment stage is

$$\int_{\mathcal{P}} C_{\mathrm{p}}(\imath)\mathrm{d}\imath + \int_{\mathcal{W}} C_{\mathrm{w}}(\imath)\mathrm{d}\imath + \int_{\mathcal{K}} C_{\mathrm{k}}(\imath)\mathrm{d}\imath = (1-p)\mathrm{q} + K,$$

the problem would become significantly more algebraically 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.

where $\mathcal{W}$ and $\mathcal{K}$ denote the sets of workers and capitalists, as determined by past investment decisions. These partition the set $\mathcal{N}$ of those outside the incumbent group.

An argument similar to Proposition 1 shows that the equilibrium rules would equalize continuation payoffs for all individuals outside the incumbent group ($\imath \in \mathcal{N}$). This means that any notional claims to capital will be set aside and individuals' payoffs will be determined according to their power, with capital redistributed accordingly under a new set of rules. Let $C_{\mathrm{n}} = C_{\mathrm{w}}(\imath) = C_{\mathrm{k}}(\imath)$ denote the common consumption level of all individuals outside the incumbent group. Also as in Proposition 1, payoffs of incumbents will be equalized, with $C_{\mathrm{p}} = C_{\mathrm{p}}(\imath)$ denoting this common payoff. The resource constraint implies that the incumbents' payoff is

$$\mathcal{U}_{\mathrm{p}} = C_{\mathrm{p}} = \frac{(1-p)(\mathrm{q} - C_{\mathrm{n}}) + K}{p}. \tag{3.4}$$

Using the argument of Proposition 1, the equilibrium rules following a rebellion at the post-investment stage can be characterized by maximizing the payoff of incumbents in [3.4] subject to a single no-rebellion constraint

$$\mathcal{U}_{\mathrm{n}} \le \mathcal{U}'_{\mathrm{p}}(K) - \delta \frac{p}{p'(K)},$$

where $p'(K)$ and $\mathcal{U}'_{\mathrm{p}}(K)$ are features of the equilibrium rules that would follow a further rebellion.[34] In equilibrium, p' and $\mathcal{U}'_{\mathrm{p}}$ may be functions of the total capital stock K, which is a fundamental state variable at this stage. Given payoff equalization and the resource constraint, the choice variables are power sharing p and the consumption C_{n} of those outside the incumbent group.

The equilibrium rules are found by solving this constrained maximization problem and then imposing the Markovian conditions $p(K) = p'(K)$ and $\mathcal{U}_{\mathrm{p}}(K) = \mathcal{U}'_{\mathrm{p}}(K)$. The unique equilibrium values of each variable are as follows, which are denoted by a † superscript to distinguish them from equilibrium rules at the pre-investment stage:

$$p^{\dagger} = \frac{\delta}{\mathrm{q} + 2\delta}, \quad \mathcal{U}^{\dagger}_{\mathrm{p}}(K) = \frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta} + K, \quad \text{and } \mathcal{U}^{\dagger}_{\mathrm{w}}(K) = \frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta} - \delta + K. \tag{3.5}$$

The equilibrium rules are such that everyone outside the incumbent group receives utility $\mathcal{U}^{\dagger}_{\mathrm{w}}(K)$. Equilibrium power sharing $p^{\dagger}$ is independent of K and is the same as that found in the endowment-economy model with linear utility from section 2.5.[35] The results show that were a rebellion to occur at the post-investment stage, the entire capital stock would be expropriated and equally distributed among the whole population. This is because the presence of the capital increases incentives for further rebellions.

[34] Under the parameter restrictions in [3.3], in equilibrium, all non-negativity constraints on consumption are slack, the constraint $p < 1/2$ is slack, and the participation constraint $\mathcal{U}^{\dagger}_{\mathrm{p}}(K) > \mathcal{U}^{\dagger}_{\mathrm{w}}(K)$ for the incumbent army is slack.

[35] This analytically convenient finding is owing to the linearity of utility in consumption.

3.2.2 Pre-investment stage rules

The stability condition for equilibrium means that rules must avoid rebellions at both the pre- and post-investment stages. Potential rebellions at these stages need to be considered separately owing to the change in the environment as a result of the accumulation of capital and the revelation of information after the investment stage.

Given the stability of equilibrium rules, an individual with an investment opportunity will decide to take it or not on the basis of the contingent consumption allocation specified by those rules. An *investor* is an individual who receives and takes an investment opportunity (and who will be referred to as a *capitalist* at the post-investment stage). If individual $\imath$ receives an investment opportunity with effort cost θ then he obtains utility $\mathcal{U}_{\mathrm{i}}(\imath;\theta) = C_{\mathrm{k}}(\imath) - \theta$ by taking the opportunity. If he chooses not to take it, he becomes a worker and obtains utility $\mathcal{U}_{\mathrm{w}} = C_{\mathrm{w}}$. This gives rise to an incentive compatibility constraint (conditional on the realization of θ) that must hold if individual $\imath$ is to invest:

$$C_{\mathrm{k}}(\imath) - C_{\mathrm{w}}(\imath) \geq \theta.$$

Let s denote the proportion of those receiving an investment opportunity for whom the incentive constraint is satisfied. The resource constraint is then

$$\int_{\mathcal{P}} C_{\mathrm{p}}(\imath)\mathrm{d}\imath + \int_{\mathcal{W}} C_{\mathrm{w}}(\imath)\mathrm{d}\imath + \int_{\mathcal{K}} C_{\mathrm{k}}(\imath)\mathrm{d}\imath = (1-p)\mathrm{q} + \mu\kappa s,$$

and the capital stock K is:

$$K = \mu\kappa s.$$

The expected utility of individual $\imath \in \mathcal{N}$ outside the incumbent group (who does not yet know whether he will become a worker or an investor) is

$$\mathcal{U}_{\mathrm{n}}(\imath) = (1-\alpha)C_{\mathrm{w}}(\imath) + \alpha\mathbb{E}_{\theta}\max\{C_{\mathrm{k}}(\imath) - \theta, C_{\mathrm{w}}(\imath)\}, \quad \text{where } \alpha = \frac{\mu}{1-p}. \qquad [3.6]$$

In this expression, α is the probability of any individual outside the incumbent group receiving an investment opportunity, which is the total measure of such opportunities divided by the measure of those individuals.[36]

The following proposition presents the key features of the equilibrium rules in this environment.

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

(i) Payoff equalization among all workers, and payoff equalization among all incumbents: $\mathcal{U}_{\mathrm{w}} = \mathcal{U}_{\mathrm{w}}(\imath)$, $\mathcal{U}_{\mathrm{p}} = \mathcal{U}_{\mathrm{p}}(\imath)$.

[36] Note that the parameter restrictions in [3.3] imply $\mu < 1/2$, and since $p < 1/2$, the number of individuals outside the group in power is always more than 50%, and hence more than the number of investment opportunities.

(ii) There is no loss from consumption equality among capitalists: $C_{\mathrm{k}} = C_{\mathrm{k}}(\imath)$, and all no-rebellion constraints for rebel armies with a positive measure of capitalists are slack. The binding constraint for capitalists is an incentive compatibility constraint that can be expressed as a threshold condition $\theta \leq \tilde{\theta}$. The effort cost threshold $\tilde{\theta}$ and the implied fraction $s = \mathbb{P}_{\theta}[\theta \leq \tilde{\theta}]$ of investment opportunities that are taken are given by:

$$\tilde{\theta} = C_{\mathrm{k}} - C_{\mathrm{w}}, \quad \text{and} \quad s = \frac{\tilde{\theta} - \psi}{\kappa - \psi}. \tag{3.7a}$$

(iii) All no-rebellion constraints at the pre-investment stage are slack. Two no-rebellion constraints at the post-investment stage are binding, namely for rebellions including only workers, and for rebellions including only incumbents:

$$\mathcal{U}_{\mathrm{w}} \geq \mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \delta\frac{p}{p^{\dagger}}, \quad \text{and} \; \mathcal{U}_{\mathrm{p}} \geq \mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \delta\frac{(p - p^{\dagger})}{p^{\dagger}}, \tag{3.7b}$$

or equivalently, any two linearly independent combinations of these no-rebellion constraints.

(iv) The binding incentive-compatibility constraint [3.7a] and the binding no-rebellion constraints [3.7b] imply that power sharing p must satisfy

$$p = p^{\dagger} + \frac{\mu\tilde{\theta}s}{\delta}. \tag{3.7c}$$

Incentives for investment ($s > 0$) thus require that power is not as concentrated as incumbents would like it to be, ex post ($p > p^{\dagger}$).

(v) Given the power-sharing constraint [3.7c], the payoff of an incumbent in terms of s is

$$\mathcal{U}_{\mathrm{p}} = \frac{(\mathrm{q} + \delta)^2}{\mathrm{q} + 2\delta} + \mu\left(\kappa - \left(\frac{\mathrm{q} + 2\delta}{\delta}\right)(\psi + (\kappa - \psi)s)\right)s. \tag{3.7d}$$

Proof See appendix A.2. ■

The intuition for the equilibrium payoff equalization among workers and among incumbents is the same as in Proposition 1. The second part of the proposition shows that no investor belongs to a rebel army associated with a binding no-rebellion constraint. The basic reason is that providing incentives to investors means granting them higher consumption than workers, and thus higher utility ex post once the sunk effort cost of investing has already been incurred (ex ante, the marginal investor has the same utility as a worker). The analysis of section 3.2.1 shows that the equilibrium rules following a rebellion at the post-investment stage will not respect individuals' holdings of capital prior to the rebellion. Thus, what investors stand to receive by participating in a rebellion (net of fighting costs) is no different from that of workers who rebel (their power is identical), while what they lose is superior if they currently hold capital. Accordingly, they are less willing to fight to replace the current rules. This implies the distribution of income needed to provide incentives to invest is not one that investors themselves could enforce by a credible threat to participate in rebellions now or at some future stage.

The fundamental problem lying behind the results of Proposition 2 is that the distribution of income needed to encourage investment diverges from that consistent with the distribution of power, and so there are incentives for groups to rebel against rules prescribing property rights. As usual, the no-rebellion constraint is binding for workers because incumbents gain by extracting as much as possible from them. What is novel here is that discouraging rebellion by workers is no longer sufficient for stability of the rules in the presence of investment opportunities: the incumbents must also worry about rebellion from within their own ranks. Incumbents would like to design rules encouraging investment by allowing investors to keep a large part of the capital they produce, but there is also the temptation ex post for them to participate in a rebellion that will allow for new rules permitting full expropriation. The fact that the effort cost of investment is sunk gives rise to a time-inconsistency problem, which is reflected in the threat of rebellion coming from inside as well as outside the group in power.[37]

Given this time-inconsistency problem, it might be thought impossible to sustain any investment in equilibrium because individuals cannot commit not to rebel. Since the defence of the current rules relies on those with power, a rebellion backed by all incumbents succeeds without requiring any fighting effort. Were the size p of the group in power equal to its equilibrium size $p^\dagger$ following a rebellion, those in power would be able to change the current rules through a costless 'suspension of the constitution'. New rules could be established while leaving all the current incumbents in power, essentially granting them full discretion ex post to rewrite the rules completely. However, when power is shared more broadly and thus $p > p^\dagger$, costless suspension of the constitution is not possible. The equilibrium size of the group in power after the rebellion is smaller than beforehand, so some incumbents must lose their positions. The rebellion launched by insiders is now necessarily a 'coup d'état' that shrinks the incumbent group. Conflict with those incumbents who would lose their positions of power makes this a costly course of action.

The formal analysis in Proposition 2 confirms that satisfaction of the no-rebellion constraints for workers and incumbents is equivalent to ensuring that power sharing p at the pre-investment stage is large in relation to $p^\dagger$, the equilibrium power sharing after a rebellion at the post-investment stage. The claim in [3.7c] is that credible limits on expropriation require more power sharing than what would be optimal for incumbents after investment decisions have actually been made. Furthermore, as the proportion s of investors rises, the required amount of power sharing p increases. Given the incentive constraint, a higher value of s requires that investors keep more of the capital they produce, which worsens the time-inconsistency problem for the incumbents.

The proposition shows that not only is this increase in power sharing *sufficient* for credible protection of property rights; it is also *necessary*. There is no other set of rules that can both establish credible incentives for investors and survive the power struggle. In particular, it might be

[37]The no-rebellion constraint for the incumbents places a lower bound on their payoff $\mathcal{U}_\mathrm{p}$ even though the rules are set up to maximize $\mathcal{U}_\mathrm{p}$ ex ante. The constraint then represents the absence of incentives to deviate from the initial rules through rebellion ex post.

thought possible to solve the problem by establishing different rules for the allocation of resources, essentially 'buying off' those who would otherwise rebel. But discouraging rebellion by workers would require transferring resources from incumbents to workers, while discouraging rebellion by incumbents would require transfers in the opposite direction. Transfers away from investors would of course destroy the very incentives that must be preserved. The only way to discourage rebellion simultaneously from both inside and outside the group in power is an increase in power sharing. Fundamentally, transfers are a zero-sum game, and can only redistribute disgruntlement with the current rules.[38]

Sharing power among a wider group of individuals therefore allows incumbents to act as a government committed to policies that would otherwise be time inconsistent. Even though all individuals act with discretion, overcoming the time-inconsistency problem is feasible. Sharing power thus emerges endogenously as a commitment device. It provides a solution to the classic problem of 'who will guard the guardians?': rules can be protected from those who hold power when some of them fear losing their privileged status if the rules are changed from within.

Broadly speaking, the extra individuals in power required to sustain property rights (the difference between p and $p^\dagger$) might be interpreted as a 'parliament', or an 'independent judiciary', or any other group with the power to resist attempts to change rules, coming especially from others in power. In the model, these extra individuals in power are in no way intrinsically different from other incumbents, and do not have access to any special technology directly protecting property rights.[39] Power sharing makes rules more stable because it makes it costlier for incumbents to replace the current rules with new ones — with potential differences in how resources and power are distributed. Once power is too concentrated, rules become subject to the whims of those in power, as noted by Montesquieu.

One early historical example that resonates with the finding of a connection between power sharing and commitment is provided by Malmendier (2009), who studies the Roman *societas publicanorum*. These were groups (precursors of the modern business corporation) to which the government contracted functions such as tax collection and public works. Their demise occurred with the transition from the Roman republic to the Roman empire. Why? According to Malmendier (2009), one possible explanation 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 government's commitment to its contract with the *societas publicanorum* and their property rights, presumably because changing the rules would result in some of the individuals in power coming into conflict with their peers, which would be costly. Once power was centralized, protection against expropriation was not possible any longer.

The essence of the argument is that while the model presupposes rules can be created, for them

[38]On the other hand, the notion of being in power is essentially an ability when fighting occurs to impose costs on others at a lower cost to oneself.

[39]In practice, the roles of these extra individuals in power are specific (legislative, judicial, etc), but the model suggests that an environment of stable rules relies not only on the actual functions of the additional incumbents, but also on power being distributed among a larger group in itself.

to be 'rules' in the sense that is commonly understood, they must be stable. The result here is that property rights can be a feature of the equilibrium rules only if power is shared among a larger group. The model precludes potential rebels from committing to create a new set of rules that is not an equilibrium, in particular, one which is not in their interests ex post. Once individuals have incurred the sunk effort costs of investing, those in power would like to sign a 'rebellion contract' where they agree to change the rules to expropriate capital, but bind themselves not to change the rules in respect of power sharing. However, each has an incentive to reduce the extent of power sharing (imposing the loss of status on others within the former incumbent group), so this contract could only be enforced by some exogenous higher authority. In the absence of such a thing, individuals may rebel against the existing rules, but cannot commit to what they will then do next.[40]

3.3 Equilibrium rules and efficient rules

Economic development ultimately requires rewarding the productive rather than just the strong, and for this to happen, rules must credibly protect the property rights of investors. It is an endogenous feature of the model that broader power sharing can sustain commitment, but is establishing such rules in the interests of those in power?

Proposition 2 shows that the equilibrium rules are subject to two binding no-rebellion constraints and a binding incentive constraint.[41] The rules specify power sharing p, and given payoff equalization, the consumption levels C_{p}, C_{w}, and C_{k}, of incumbents, workers, and capitalists, respectively. Using the resource constraint and the power-sharing constraint (which combines the two binding no-rebellion constraints), the payoff of incumbents can be written in terms of the fraction s of investment opportunities that are taken, which is linked to consumption levels via the incentive constraint. This payoff is given in equation [3.7d], which is maximized by the following choice of s:

$$s^* = \max\left\{0, \frac{\delta\kappa - (\mathrm{q} + 2\delta)\psi}{2(\mathrm{q} + 2\delta)(\kappa - \psi)}\right\}. \qquad [3.8]$$

As confirmed below in Proposition 3, this characterizes the unique equilibrium of the model. But does it correspond to the efficient level of investment?

To study the question of efficiency, the payoffs of all individuals are derived given the measure p of incumbents and the fraction s of investment opportunities that are taken. The values of p and s imply that there are $1 - p - \mu s$ workers, and μs investors with utility $\mathcal{U}_{\mathrm{i}}(\theta) = C_{\mathrm{k}} - \theta$ (those for whom the realization of θ is no more than $\tilde{\theta}$). The average utility $\bar{\mathcal{U}}$ over all these groups is

$$\bar{\mathcal{U}} \equiv \int_{\Omega} \mathcal{U}(\imath)\mathrm{d}\imath = p\mathcal{U}_{\mathrm{p}} + (1 - p - \mu s)\mathcal{U}_{\mathrm{w}} + \mu s \mathbb{E}_{\theta}[\mathcal{U}_{\mathrm{i}}(\theta) | \theta \leq \tilde{\theta}].$$

[40] If it were possible for rebels to commit to restrict reoptimization to certain areas following a rebellion then paradoxically this makes it harder to establish rules that allow for credible commitments. For example, suppose the distribution of power is defined on the first page of the constitution, and limits on expropriation of private property are specified on the second page. If it were feasible somehow to prevent a successful rebellion from rewriting page one of the constitution then this would annihilate the credibility of page two.

[41] It is verified in Proposition 3 that the non-negativity constraint $C_{\mathrm{w}} \geq 0$, the size constraint $p < 1/2$, and the incumbent-army participation constraints $\mathcal{U}_{\mathrm{p}} > \mathcal{U}'_{\mathrm{w}}$ and $\mathcal{U}_{\mathrm{p}} > \mathcal{U}^{\dagger}_{\mathrm{w}}(K)$ are all slack in equilibrium.

The average utility can be written in terms of the expected surplus $\mathcal{S}_{\mathrm{i}}(\tilde{\theta})$ from receiving an investment opportunity:

$$\bar{\mathcal{U}} = p\mathcal{U}_{\mathrm{p}} + (1-p)\mathcal{U}_{\mathrm{w}} + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}), \quad \text{where } \mathcal{S}_{\mathrm{i}}(\tilde{\theta}) \equiv \mathbb{E}_{\theta}\max\{\tilde{\theta}-\theta, 0\}. \tag{3.9}$$

Since θ lies in the range $[\psi, \kappa]$, it is easy to see that the surplus $\mathcal{S}_{\mathrm{i}}(\tilde{\theta})$ from investment opportunities is maximized when $\hat{\tilde{\theta}} = \kappa$, that is, when all investment opportunities are taken ($s = 1$). With no restrictions on the distribution of resources, this choice maximizes $\bar{\mathcal{U}}$, and is hence the first-best level of investment. However, the first best is not the most interesting welfare benchmark. The model demonstrates that power sharing is required for protection of property rights, but a larger group in power diverts more individuals from directly productive occupations (individuals in power do not receive the endowment q). This means that power sharing entails an opportunity cost.

Taking that opportunity cost into account, consider the following notion of *constrained efficiency*. Suppose that it were possible exogenously to impose some level of investment (and thus the variable s) at all stages when rules are chosen. All other aspects of the rules are determined endogenously as before subject to ensuring s is consistent with investors' incentive compatibility constraints and the no-rebellion constraints arising from the power struggle. The constrained-efficient level of s is what would then be chosen by a benevolent agent who takes into account the environment in which equilibrium rules are determined. The benevolent agent would appreciate that more investment requires greater protection of property rights, and thus more power sharing. The concept of constrained efficiency requires setting the benefit of more investment against the resource cost of power sharing.[42] In the public-good example considered in section 2.6, a benevolent agent could not improve upon the efficiency of the rules by imposing a level of public-good provision different from what prevails in equilibrium. Here, the issue is whether the equilibrium amount of investment coincides with its constrained-efficient level.

To find the constrained-efficient level of investment, the benevolent agent maximizes the average utility $\bar{\mathcal{U}}$ of all individuals subject to the binding constraints identified in Proposition 2 (the configuration of binding constraints is the same even with an exogenously given level of $s > 0$). As a consequence of [3.7b], workers' and incumbents' payoffs are tied together by $\mathcal{U}_{\mathrm{w}} = \mathcal{U}_{\mathrm{p}} - \delta$. Since the benevolent agent takes such constraints into account, this relationship is substituted into the expression for $\bar{\mathcal{U}}$:

$$\bar{\mathcal{U}} = \mathcal{U}_{\mathrm{p}} - \delta(1-p) + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}). \tag{3.10}$$

There are thus two differences between the expressions for average utility $\bar{\mathcal{U}}$ and the incumbents' utility $\mathcal{U}_{\mathrm{p}}$. The second term on the right-hand side is related to the distribution of resources between individuals with different levels of power, and the third term reflects the investors' surplus.

The expression for $\bar{\mathcal{U}}$ is maximized by noting that the binding constraints imply $\mathcal{U}_{\mathrm{p}}$ is given by

[42] If there were no resource cost of increasing the size of the group in power then the notion of constrained efficiency would coincide with first best. Note that in the model with the public-good technology from section 2.6, the first-best and constrained-efficient outcomes are the same.

[3.7d], and that these constraints are equivalent to the power-sharing constraint [3.7c] linking p and s, and by using the expression for the investors' surplus from [3.9].

Proposition 3 *(i) The unique equilibrium s^* is given by the expression in [3.8].*

(ii) The constrained-efficient level of s (denoted by $s^\diamond$), which maximizes $\bar{\mathcal{U}}$ in [3.10], is given by the following expression (if the non-negativity constraint on workers' consumption does not bind):

$$s^\diamond = \max\left\{0, \frac{\delta\kappa - (q+\delta)\psi}{(2q+\delta)(\kappa-\psi)}\right\}. \qquad [3.11]$$

(iii) s^* is positive when $\kappa/\psi - 1 > 1 + q/\delta$, while $\kappa/\psi - 1 > q/\delta$ is necessary for $s^\diamond > 0$. Whenever $s^\diamond > 0$, it must be the case that $s^* < s^\diamond$.

Proof See appendix A.3. ■

There are two reasons why the constrained-efficient level of investment is larger than what prevails in equilibrium. The first (and more interesting) distortion follows from the distributional consequences of protection against expropriation (the second term on the right-hand side of equation [3.10]). Commitment to property rights requires sharing power, which in turn requires sharing rents because incumbents are more powerful than other individuals. Incumbents have access to an endogenous commitment device through power sharing that can implement the constrained-efficient outcome, but the need to avoid rebellions means this entails sharing rents. Therefore, the cost to incumbents of expanding their number is not simply the lost output from diverting individuals away from directly productive activities.

In contrast to the public-good example of section 2.6, in an environment with investment, the power struggle imposes an endogenous and binding constraint on the set of possible transfers among individuals. This leads to a breakdown of the political analogue of the Coase theorem. Power sharing gives rise to commitment, but the association between power and rents places a lower bound on the consumption of each individual in power. The impossibility of sharing power without sharing rents thus drives a wedge between maximizing total output and maximizing an incumbent's payoff.[43]

The second distortion that results in equilibrium investment being too low is that the rules determined in equilibrium do not take account of investors' surplus (the third term on the right-hand side of equation [3.10]). Since investors' effort costs θ are not public information, it is impossible for rules to specify a consumption allocation contingent on this information. The no-rebellion constraints for rebel armies including non-marginal investors will therefore be slack, so no benefit

[43]The welfare implications of the power parameter δ are non-trivial in this application of the model. In the public-good example of section 2.6, a larger δ can only be harmful to workers because it allows more rents to be extracted, resulting in a more unequal distribution of income. In contrast, in an economy with a very small value of δ, there would not be any investment in equilibrium. A larger δ parameter makes it easier for incumbents to remain in power, which directly benefits them, but might also allow them to offer some protection of property rights.

accrues to incumbents from increases in such investors' payoffs. This increases the wedge between total output and the payoff of incumbents.[44]

Acemoglu, Johnson and Robinson (2005) present evidence that institutional failures in providing adequate protection of property right are especially damaging to economic performance. But why should property rights be so susceptible to political failures compared to other aspects of institutions? The model here sheds light on this question by explaining why there is often tenacious opposition by incumbents to institutions consistent with credible property rights. For example, in seventeenth-century England, the Glorious Revolution led to power sharing between king and parliament. By accepting the Bill of Rights, King William III conceded that power would be shared. North and Weingast (1989) argue that the Glorious Revolution began an era of secure property rights and put an end to confiscatory government. As a result, the English government was able to borrow much more, and at substantially lower rates. This was certainly in the interests of the king, yet the earlier Stuart kings had staunchly resisted sharing power with parliament. According to the model, secure property rights require just such power sharing to make it 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 total pie becomes larger, with a smaller share, the amount received by the king might end up being lower.

4 Concluding remarks

Research in economics has frequently progressed by focusing on the behaviour of individuals subject to some fundamental constraints or frictions and deriving the resulting implications for the economy. For example, it is often claimed that unemployment, credit rationing, and missing markets ought not to be directly assumed, but instead derived from the likes of search frictions, limited pledgeability, or asymmetric information. This paper proposes a model of the rules that emerge from the power struggle which we think of as an attempt to model politics in that tradition. The building blocks of the model are the basics of preferences, technologies, and a single rebellion mechanism that allows individuals to form groups and fight for power. Those in power establish rules and have an advantage in defending the status quo, but the option of rebelling against the rules is open to everyone on the same terms.

The model is used to study an environment where investment is possible but can be expropriated. In order to generate commitment to rules that would otherwise be time inconsistent, the incumbent group is endogenously enlarged. But the same conflict mechanism that explains how power sharing can overcome the commitment problem also implies that sharing power cannot be done without sharing rents. This imposes endogenous limits on the set of possible allocations and precludes Pareto-improving deals. In equilibrium, there is too little power sharing, and thus not enough commitment to offer investors the strong property rights required for economic efficiency.

[44] The first and second distortions correspond respectively to the second and third terms in the expression for $\bar{\mathcal{U}}$ in [3.10]. The effects of each of them on the first-order condition determining s^* are thus seen to operate independently of the other distortion.

This paper is voluntarily abstracting from several real-world features and frictions, but the model could be extended in order to study other questions. The remainder of this section explores some directions for future research.

The model takes the fighting strength of a member of the incumbent army as given, but in some applications it could be important to consider ways of increasing the cost of rebellions other than having a larger incumbent group. Campante, Do and Guimaraes (2013) extend the framework developed here to build a model where the cost of rebellion depends on how far a citizen is from the capital city. This means that isolating the capital city raises the cost of rebellions. The model yields predictions on the relationship between governance and the isolation of the capital city which are consistent with the data. Increasing military spending is another way the cost of rebellion might be raised. This is also analysed in Campante, Do and Guimaraes (2013), and the finding is that military spending and isolating the capital city work as substitutes.

Guimaraes and Sheedy (2013) extend the framework to a world of open economies where the possibility of international trade affects the relative price of two goods. For the reasons highlighted here, production of an investment good requires power sharing because it depends on protection of property rights, while the other good is modelled as an endowment. The model predicts that trade generates divergence as production of the investment good becomes concentrated in a subset of economies where power is shared broadly, so trade benefits some economies, but harms others.

Beyond this existing work, it would also be possible to introduce a stochastic element into the model to allow rebellions to occur in equilibrium. This would make the probability of rebellion an important consideration for those in power. Furthermore, an extension of the model where rebellions do not wipe the slate completely clean could be used to study the evolution of institutions. Many other interesting extensions remain.

References

Acemoglu, D. (2003), "Why not a political Coase theorem? Social conflict, commitment, and politics", *Journal of Comparative Economics*, **31(4)**:620–652. 5, 16

——— (2008), "Oligarchic versus democratic societies", *Journal of the European Economic Association*, **6(1)**:1–44 (March). 5

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

——— (2012), "Dynamics and stability of constitutions, coalitions, and clubs", *American Economic Review*, **102(4)**:1446–1476 (June). 4

Acemoglu, D. and Johnson, S. (2005), "Unbundling institutions", *Journal of Political Economy*, **113(5)**:949–995. 1

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. 5, 28

Acemoglu, D. and Robinson, J. A. (2000), "Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective", *Quarterly Journal of Economics*, **115(4)**:1167–1199 (November). 4

——— (2006), *Economic Origins of Dictatorship and Democracy*, Cambridge University Press. 5

——— (2008), "Persistence of power, elites, and institutions", *American Economic Review*, **98(1)**:267–293 (March). 5

Bai, J. H. and Lagunoff, R. (2011), "On the Faustian dynamics of policy and political power", *Review of Economic Studies*, **78(1)**:17–48. 4

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

Basu, K. (2000), *Prelude to Political Economy*, Oxford University Press. 4

Besley, T. and Persson, T. (2009a), "The origins of state capacity: Property rights, taxation, and politics", *American Economic Review*, **99(4)**:1218–1244 (September). 5

——— (2009b), "Repression or civil war?", *American Economic Review*, **99(2)**:292–297 (May). 5

——— (2010), "State capacity, conflict and development", *Econometrica*, forthcoming. 5

Campante, F., Do, Q.-A. and Guimaraes, B. (2013), "Isolated capital cities and misgovernance: Theory and evidence", working paper, Kennedy School of Government. 29

Engerman, S. L. and Sokoloff, K. L. (1997), "Factor endowments, institutions, and differential paths of growth among New World economies: 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. 260–304. 5

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. 4

Glaeser, E., Scheinkman, J. and Shleifer, A. (2003), "The injustice of inequality", *Journal of Monetary Economics*, **50(1)**:199–222 (January). 5

Greif, A. (2006), *Institutions and the Path to the Modern Economy*, Cambridge University Press. 5

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). 4

Guimaraes, B. and Sheedy, K. D. (2013), "Trade, power sharing, and divergence", working paper. 29

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). 5

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). 5

Hirshleifer, J. (1995), "Anarchy and its breakdown", *Journal of Political Economy*, **103(1)**:26–52 (February). 4

Jack, W. and Lagunoff, R. (2006), "Dynamic enfranchisement", *Journal of Public Economics*, **90(4–5)**:551–572 (May). 4

Levy, G. (2004), "A model of political parties", *Journal of Economic Theory*, **115(2)**:250–277 (April). 4

Mailath, G., Morris, S. and Postlewaite, A. (2001), "Laws and authority", working paper, Princeton University. 4

Malmendier, U. (2009), "Law and finance at the origin", *Journal of Economic Literature*, **47(4)**:1076–1108. 24

Myerson, R. B. (1993), "Incentives to cultivate favored minorities under alternative electoral rules", *American Political Science Review*, **87(4)**:856–869 (December). 9

——— (2004), "Justice, institutions and multiple equilibria", *Chicago Journal of International Law*, **5(1)**:91–107. 4

——— (2008), "The autocrat's credibility problem and foundations of the constitutional state", *American Political Science Review*, **102(1)**:125–139. 13

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

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

North, D. C. (1990), *Institutions, Institutional Change, and Economic Performance*, Cambridge University Press. 5

North, D. C. and Weingast, B. R. (1989), "Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England", *Journal of Economic History*, **49(4)**:803–832 (December). 5, 28

Persson, T. and Tabellini, G. (2000), *Political Economics: Explaining Economic Policy*, MIT Press. 17

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

Ray, D. (2007), *A Game-Theoretic Perspective on Coalition Formation*, Oxford University Press. 4

Schelling, T. C. (1960), *The Strategy of Conflict*, Harvard University Press. 4

A Technical appendix

A.1 Proof of Proposition 1

There must be a binding no-rebellion constraint for a rebel army including a positive measure of workers

If not, there must be a positive measure of workers for whom inclusion in a rebel army (displacing incumbents) strictly reduces the effective fighting strength of the rebel army (effective fighting strength means the fighting effort of the army's members plus any loss of δ for the incumbent army when some incumbents join the rebel army). It follows that consumption of these workers can be reduced by some positive amount, allowing the average incumbent payoff to be increased, but without implying that any rebellion would now be successful.

Payoff equalization among workers

Since the size of any incumbent group cannot exceed the number of workers, a rebel army never includes all workers. Now suppose there is payoff inequality among workers. It is known that a binding no-rebellion constraint must apply to a rebel army with a positive measure of workers. Consider a deviation where consumption of all workers is set equal to the previous average level of workers' consumption. Given weakly concave utility, this is feasible, and with strictly concave utility, it leaves surplus resources that can now be distributed among the incumbents. Since the previous strongest rebel army including workers would include those with the lowest payoffs, following this deviation, the maximum fighting strengths of all rebel armies including workers are reduced. This slackens a binding constraint, so payoff inequality among workers is inconsistent with equilibrium.

Power determines rents in equilibrium

Let $\bar{\mathcal{U}}_{\mathrm{p}}^*$ denote the average incumbent payoff in equilibrium. The claim is that $\bar{\mathcal{U}}_{\mathrm{p}}^* = \mathcal{U}_{\mathrm{w}}^* + \delta$. To prove this, first suppose $\bar{\mathcal{U}}_{\mathrm{p}}^* > \mathcal{U}_{\mathrm{w}}^* + \delta$. Since $\mathcal{U}_{\mathrm{p}}' = \bar{\mathcal{U}}_{\mathrm{p}}^*$ in equilibrium, it would follow that a rebel army including only workers (of which there are enough) would put in fighting effort strictly greater than δp^* (given that the rebel army would have size $p' = p^*$ in equilibrium). This rebellion would succeed, hence it must be that $\bar{\mathcal{U}}_{\mathrm{p}}^* \leq \mathcal{U}_{\mathrm{w}}^* + \delta$.

Now suppose $\bar{\mathcal{U}}_{\mathrm{p}}^* < \mathcal{U}_{\mathrm{w}}^* + \delta$. Consider the following deviation. First, redistribute incumbents' consumption equally among them. Given weakly concave utility, the new common payoff $\mathcal{U}_{\mathrm{p}}$ is at least as large as $\bar{\mathcal{U}}_{\mathrm{p}}^*$. Next, reduce consumption of all workers by a positive amount, but small enough so that $\bar{\mathcal{U}}_{\mathrm{p}}^* \leq \mathcal{U}_{\mathrm{w}} + \delta$ holds, where $\mathcal{U}_{\mathrm{w}}$ is the new payoff of workers. The reduction in workers' consumption allows the common incumbent payoff to be increased, ensuring $\mathcal{U}_{\mathrm{p}} > \bar{\mathcal{U}}_{\mathrm{p}}^*$. Given that this means $\mathcal{U}_{\mathrm{p}} > \mathcal{U}_{\mathrm{p}}'$, no-one in the group in power is now willing to join a rebel army. A rebel army can then only comprise workers, but the deviation is constructed to ensure such a rebellion would not succeed. The claim is proved.

Payoff equalization among incumbents

Given that $\bar{\mathcal{U}}_{\mathrm{p}}^* = \mathcal{U}_{\mathrm{w}}^* + \delta$ and $\mathcal{U}_{\mathrm{p}}' = \bar{\mathcal{U}}_{\mathrm{p}}^*$ hold in equilibrium, it follows that each worker is willing to put in exactly δ fighting effort in a rebel army. Now suppose there is payoff inequality among a positive measure of incumbents, which means that a positive measure d of incumbents receive payoffs strictly below the average $\bar{\mathcal{U}}_{\mathrm{p}}^*$. These are willing to rebel and to put in strictly positive fighting effort. A rebel army including these incumbents and with the remaining places filled by a measure $p^* - d$ of workers would have total fighting strength larger than $\delta(p^* - d)$, which is the fighting strength of the incumbent army. Therefore, payoff inequality among incumbents is inconsistent with equilibrium. The common incumbent payoff must then satisfy $\mathcal{U}_{\mathrm{p}}^* - \mathcal{U}_{\mathrm{w}}^* = \delta$.

Equivalence to a single no-rebellion constraint

With payoff equalization among all workers and among all incumbents, the choice of rules reduces to a choice of power sharing p and a common level of consumption for workers C_{w} (with consumption of incumbents determined by the resource constraint). Equilibrium rules are then the solution of the following maximization problem

$$\max_{p, C_{\mathrm{w}}} \mathcal{U}_{\mathrm{p}} \text{ subject to } \sigma(\mathcal{U}_{\mathrm{p}}' - \mathcal{U}_{\mathrm{w}}) + (1-\sigma)\mathbb{1}[\mathcal{U}_{\mathrm{p}} \leq \mathcal{U}_{\mathrm{p}}'](\mathcal{U}_{\mathrm{p}}' - \mathcal{U}_{\mathrm{p}} + \delta) \leq \delta \frac{p}{p'} \text{ for all } \sigma \in \left[\max\left\{0, \frac{p'-p}{p'}\right\}, 1\right], \quad \text{[A.1.1]}$$

where σ denotes the fraction of places in the post-rebellion incumbent group assigned to those who are currently workers. An equilibrium is characterized by the values of p^* and C_w^* that solve the maximization problem [A.1.1] taking p' and $\mathcal{U}_p'$ as given, but with $p' = p^*$ and $\mathcal{U}_p' = \mathcal{U}_p^*$ in equilibrium.

Now consider a simpler maximization problem where the no-rebellion constraint in [A.1.1] is imposed only for $\sigma = 1$:

$$\max_{p, C_w} \mathcal{U}_p \text{ subject to } \mathcal{U}_p' - \mathcal{U}_w \leq \delta \frac{p}{p'}. \qquad \text{[A.1.2]}$$

An equilibrium of this alternative problem is defined as a solution $(p^\star, C_w^\star)$ of the maximization problem [A.1.2], taking p' and $\mathcal{U}_p'$ as given, but with $p' = p^\star$ and $\mathcal{U}_p' = \mathcal{U}_p^\star$ in equilibrium.

Start by considering an equilibrium $(p^\star, C_w^\star)$ of the simpler problem [A.1.2]. Since these rules must satisfy the no-rebellion constraint from [A.1.2] and $p' = p^\star$ and $\mathcal{U}_p' = \mathcal{U}_p^\star$, it follows that $\mathcal{U}_p^\star - \mathcal{U}_w^\star \leq \delta$. Therefore, for any $\sigma \in [0, 1]$:

$$\sigma(\mathcal{U}_p^\star - \mathcal{U}_w^\star) + (1-\sigma)\mathbb{1}[\mathcal{U}_p^\star \leq \mathcal{U}_p^\star](\mathcal{U}_p^\star - \mathcal{U}_p^\star + \delta) = \sigma(\mathcal{U}_p^\star - \mathcal{U}_w^\star) + (1-\sigma)\delta \leq \sigma\delta + (1-\sigma)\delta = \delta = \delta\frac{p^\star}{p^\star}.$$

This demonstrates that $(p^\star, C_w^\star)$ satisfies the no-rebellion constraints of the original problem [A.1.1] when $p' = p^\star$ and $\mathcal{U}_p' = \mathcal{U}_p^\star$. Now take any other rules (p, C_w) not subject to rebellion in the original problem [A.1.1], again when $p' = p^\star$ and $\mathcal{U}_p' = \mathcal{U}_p^\star$. Evaluating the no-rebellion constraint at $\sigma = 1$ yields $\mathcal{U}_p^\star - \mathcal{U}_w \leq \delta p/p^\star$, which shows that these rules also satisfy the no-rebellion constraint of the simpler problem [A.1.2]. Since $(p^\star, C_w^\star)$ maximizes $\mathcal{U}_p$ over all rules satisfying the constraint in [A.1.2], it must be the case that $\mathcal{U}_p \leq \mathcal{U}_p^\star$ for any rules (p, C_w) consistent with the constraint in [A.1.1]. Therefore, $(p^\star, C_w^\star)$ is also an equilibrium of [A.1.1] as well.

Now consider the converse. Take an equilibrium (p^*, C_w^*) of the original problem [A.1.1]. These rules are clearly consistent with the constraint in [A.1.2] when $p' = p^*$ and $\mathcal{U}_p' = \mathcal{U}_p^*$ because the constraint is a special case of that in [A.1.1] when $\sigma = 1$. Suppose for contradiction that (p^*, C_w^*) is not an equilibrium of the problem [A.1.2]. Since it satisfies the no-rebellion constraint, it must therefore be the case that there exists another set of rules (p, C_w) such that $\mathcal{U}_p > \mathcal{U}_p^*$ and satisfying the no-rebellion constraint in [A.1.2] when $p' = p^*$ and $\mathcal{U}_p' = \mathcal{U}_p^*$. Now take any $\sigma \in [0, 1]$ and multiply both sides of the inequality in [A.1.2] by this number to obtain:

$$\sigma(\mathcal{U}_p^* - \mathcal{U}_w) \leq \sigma\delta\frac{p}{p^*} \leq \delta\frac{p}{p^*}.$$

Observe that $(1-\sigma)\mathbb{1}[\mathcal{U}_p \leq \mathcal{U}_p^*](\mathcal{U}_p^* - \mathcal{U}_p + \delta) = 0$, so this demonstrates that (p, C_w) satisfies [A.1.1] for all $\sigma \in [\underline{\sigma}, 1]$. Since these rules satisfy the no-rebellion constraint in [A.1.1], the resulting incumbent payoff cannot be higher than the incumbent payoff in equilibrium, hence $\mathcal{U}_p \leq \mathcal{U}_p^*$. This contradicts the inequality $\mathcal{U}_p > \mathcal{U}_p^*$ obtained earlier, and thus proves that (p^*, C_w^*) must be an equilibrium of the simpler problem [A.1.2].

In summary, it has been shown that the set of equilibria of the original problem [A.1.1] is identical to the set of equilibria of the simpler problem [A.1.2]. Therefore, there is no loss of generality in imposing [2.6] as the only no-rebellion constraint.

Existence and uniqueness of the equilibrium

With payoff equalization, the resource constraint implies $C_p = (1-p)(q - C_w)/p$. Worker and incumbent payoffs are given by $\mathcal{U}_w = u(C_w)$ and $\mathcal{U}_p = u((1-p)(q - C_w)/p)$ respectively. It has been shown that the equilibrium rules can be characterized as a solution of [A.1.2]. Any equilibrium (p^*, C_w^*) is thus a solution of the maximization problem

$$\max_{p, C_w} u\left(\frac{(1-p)(q - C_w)}{p}\right) \text{ subject to } \mathcal{U}_p^* - u(C_w) \leq \delta\frac{p}{p^*}, \qquad \text{[A.1.3]}$$

taking p^* and $\mathcal{U}_p^*$ as given, but with $p = p^*$ and $\mathcal{U}_p = \mathcal{U}_p^*$ in equilibrium. The solution of the maximization problem must also respect the constraint $p < 1/2$, and the non-negativity constraints on all individuals' consumption, which are equivalent to $0 \leq C_{\mathrm{w}} \leq \mathrm{q}$ here.

It has already been seen that the no-rebellion constraint must be binding, which was reduced to a single equation in [A.1.3]. This equation can be solved to obtain C_{w} as a function of p for given values of p^* and $\mathcal{U}_{\mathrm{p}}^*$:

$$C_{\mathrm{w}} = u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta \frac{p}{p^*}\right). \qquad [A.1.4]$$

Differentiating shows that the constraint implicitly specifies a negative relationship between C_{w} and p:

$$\frac{\partial C_{\mathrm{w}}}{\partial p} = -\frac{\delta}{p^*}\frac{1}{u'\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)} < 0. \qquad [A.1.5]$$

The problem [A.1.3] is equivalent to maximizing $C_{\mathrm{p}} = (1-p)(q - C_{\mathrm{w}})/p$ over values of p after substituting for C_{w} using equation [A.1.4]:

$$\max_p \frac{(1-p)}{p}\left(\mathrm{q} - u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right), \qquad [A.1.6]$$

subject to $p < 1/2$ and the value of C_{w} implied by [A.1.4] being such that $0 \leq C_{\mathrm{w}} \leq \mathrm{q}$. Taking the derivative of C_{p} with respect to p (making use of [A.1.5]):

$$\frac{\partial C_{\mathrm{p}}}{\partial p} = \frac{1}{p^2}\left(\frac{p}{p^*}\frac{\delta(1-p)}{u'\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)} - \left(\mathrm{q} - u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)\right). \qquad [A.1.7]$$

The second derivative is

$$\frac{\partial^2 C_{\mathrm{p}}}{\partial p^2} = -\frac{2}{u'\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)}\frac{\delta p}{p^*} + \frac{u''\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)}{\left\{u'\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\frac{p}{p^*}\right)\right)\right\}^3} - \frac{2}{p}\frac{\partial C_{\mathrm{p}}}{\partial p}. \qquad [A.1.8]$$

In equilibrium it is necessary to have $p = p^*$, in which case the binding no-rebellion constraint [A.1.4] reduces to

$$C_{\mathrm{w}}^* = u^{-1}(\mathcal{U}_{\mathrm{p}}^* - \delta),$$

or equivalently

$$u\left(\frac{(1-p^*)(\mathrm{q} - C_{\mathrm{w}}^*)}{p^*}\right) = \mathcal{U}_{\mathrm{p}}^* = u(C_{\mathrm{w}}^*) + \delta. \qquad [A.1.9]$$

This equation is in turn equivalent to

$$p^* u^{-1}(u(C_{\mathrm{w}}^*) + \delta) = (1-p^*)(\mathrm{q} - C_{\mathrm{w}}^*). \qquad [A.1.10]$$

Evaluating the derivative of C_{p} from [A.1.7] at $p = p^*$:

$$\left.\frac{\partial C_{\mathrm{p}}}{\partial p}\right|_{p=p^*} = \frac{1}{p^{*2}}\left(\frac{\delta(1-p^*)}{u'\left(u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\right)\right)} - \left(\mathrm{q} - u^{-1}\left(\mathcal{U}_{\mathrm{p}}^* - \delta\right)\right)\right).$$

Since [A.1.9] implies that $u^{-1}(\mathcal{U}_{\mathrm{p}}^* - \delta) = C_{\mathrm{w}}^*$, this can be simplified as follows:

$$\left.\frac{\partial C_{\mathrm{p}}}{\partial p}\right|_{p=p^*} = \frac{1}{p^{*2}}\left(\frac{\delta(1-p^*)}{u'\left(C_{\mathrm{w}}^*\right)} - (\mathrm{q} - C_{\mathrm{w}}^*)\right). \qquad [A.1.11]$$

Now define the following functions $\mathscr{G}(p, \varkappa)$ and $\mathscr{U}(p, \varkappa)$, which will represent respectively (with $\varkappa = C_{\mathrm{w}}$) the no-rebellion constraint and the first-order condition in equilibrium:

$$\mathscr{G}(p, \varkappa) \equiv p u^{-1}(u(\varkappa) + \delta) - (1-p)(\mathrm{q} - \varkappa), \quad \mathscr{U}(p, \varkappa) \equiv \delta(1-p) - (\mathrm{q} - \varkappa)u'(\varkappa). \tag{A.1.12}$$

The no-rebellion constraint [A.1.10] for an equilibrium holds when $\mathscr{G}(p^*, C_{\mathrm{w}}^*)$ is zero, while the derivative of C_{p} in [A.1.11] is proportional to $\mathscr{U}(p^*, C_{\mathrm{w}}^*)$:

$$\mathscr{G}(p^*, C_{\mathrm{w}}^*) = 0, \quad \text{and} \quad \left.\frac{\partial C_{\mathrm{p}}}{\partial p}\right|_{p=p^*} = \frac{1}{p^{*2}} \frac{1}{u'(C_{\mathrm{w}}^*)} \mathscr{U}(p^*, C_{\mathrm{w}}^*). \tag{A.1.13}$$

The partial derivatives of the function $\mathscr{G}(p, \varkappa)$ from [A.1.12] are:

$$\frac{\partial \mathscr{G}}{\partial p} = u^{-1}(u(\varkappa) + \delta) + (\mathrm{q} - \varkappa), \quad \frac{\partial \mathscr{G}}{\partial \varkappa} = p \frac{u'(\varkappa)}{u'\left(u^{-1}(u(\varkappa) + \delta)\right)} + (1-p). \tag{A.1.14}$$

Since $u(\cdot)$ is strictly increasing, $u'(\varkappa) > 0$ and $u'(u^{-1}(u(\varkappa) + \delta)) > 0$, and also $u^{-1}(u(\varkappa) + \delta) > \varkappa$. It follows that both of the partial derivatives above are strictly positive for all $0 \leq p < 1/2$ and $0 \leq \varkappa \leq \mathrm{q}$. The partial derivatives of the function $\mathscr{U}(p, \varkappa)$ from [A.1.12] are:

$$\frac{\partial \mathscr{U}}{\partial p} = -\delta, \quad \frac{\partial \mathscr{U}}{\partial \varkappa} = u'(\varkappa) - (\mathrm{q} - \varkappa)u''(\varkappa). \tag{A.1.15}$$

The properties of $u(\cdot)$ ensure that $u'(\varkappa) > 0$ and $u''(\varkappa) \leq 0$, so $\mathscr{U}(p, \varkappa)$ is strictly decreasing in p and strictly increasing in $\varkappa$.

Now consider two functions $\mathscr{H}(p)$ and $\mathscr{V}(p)$ defined implicitly by the equations:

$$\mathscr{G}(p, \mathscr{H}(p)) = 0, \quad \text{and} \quad \mathscr{U}(p, \mathscr{V}(p)) = 0. \tag{A.1.16}$$

Where these functions are defined, the signs of their partial derivatives can be deduced using [A.1.14] and [A.1.15]:

$$\mathscr{H}'(p) = -\frac{\partial \mathscr{G}}{\partial p} \Big/ \frac{\partial \mathscr{G}}{\partial \varkappa} < 0, \quad \text{and} \quad \mathscr{V}'(p) = -\frac{\partial \mathscr{U}}{\partial p} \Big/ \frac{\partial \mathscr{U}}{\partial \varkappa} > 0. \tag{A.1.17}$$

Observe from [A.1.12] that $\mathscr{U}(1, \varkappa) = -(\mathrm{q} - \varkappa)u'(\varkappa)$, with the definition [A.1.16] then implying $\mathscr{V}(1) = \mathrm{q}$ given $u'(\cdot) > 0$. [A.1.17] shows that $\mathscr{V}(p)$ is strictly increasing in p, so it follows by continuity that either there exists a $\underline{\pi} > 0$ such that $\mathscr{V}(\underline{\pi}) = 0$, or $\mathscr{V}(0) \geq 0$, in which case $\underline{\pi}$ is set to zero. With the resulting $\underline{\pi} \in [0, 1)$, define $\underline{\varkappa} \equiv \mathscr{V}(\underline{\pi})$, noting that this satisfies $0 \leq \underline{\varkappa} < \mathrm{q}$ because $\mathscr{V}(\underline{\pi}) \geq 0$ and $\mathscr{V}(1) = \mathrm{q}$. The function $\mathscr{V}(p)$ is then well defined on the interval $[\underline{\pi}, 1]$ in the sense of returning a value of $\varkappa$ in the interval $[\underline{\varkappa}, \mathrm{q}]$.

It can be seen from [A.1.12] that $\mathscr{G}(0, \varkappa) = -(\mathrm{q} - \varkappa)$, so the definition in [A.1.16] implies $\mathscr{H}(0) = \mathrm{q}$. Note that since the utility function $u(\cdot)$ is strictly increasing, so is its inverse $u^{-1}(\cdot)$. It follows that $u^{-1}(u(\varkappa) + \delta) > \varkappa$ and thus for all $p > 0$:

$$\mathscr{G}(p, \varkappa) > p\varkappa - (1-p)(\mathrm{q} - \varkappa) = \varkappa - \mathrm{q}(1-p).$$

Using [A.1.16], this means that $0 = \mathscr{G}(p, \mathscr{H}(p)) > \mathscr{H}(p) - \mathrm{q}(1-p)$, and hence

$$\mathscr{H}(p) < \mathrm{q}(1-p), \quad \text{for all } \; p > 0.$$

[A.1.17] shows that $\mathscr{H}(p)$ is strictly decreasing in p, so given $\mathscr{H}(0) = \mathrm{q}$ and the bound above, it follows by continuity that there exists a $\overline{\pi} \in (0, 1)$ such that $\mathscr{H}(\overline{\pi}) = 0$. The function $\mathscr{H}(p)$ is then well defined on the interval $[0, \overline{\pi}]$ in the sense of returning a value of $\varkappa$ in the interval $[0, \mathrm{q}]$.

Let $\mathscr{H}^{-1}(\varkappa)$ denote the inverse function of $\mathscr{H}(p)$, defined on $[0, \mathrm{q}]$. Similarly, $\mathscr{V}^{-1}(\varkappa)$ is the inverse function of $\mathscr{V}(p)$, defined on $[\underline{\varkappa}, \mathrm{q}]$, where $0 \leq \underline{\varkappa} \equiv \mathscr{V}(\underline{\pi}) < \mathrm{q}$. Since $\mathscr{H}(p)$ is strictly decreasing and $\mathscr{V}(p)$

is strictly increasing according to [A.1.17], their inverse functions inherit these properties. Now define the following function $\mathscr{A}(\varkappa)$ on $[\underline{\varkappa}, \mathrm{q}]$:

$$\mathscr{A}(\varkappa) \equiv \mathscr{V}^{-1}(\varkappa) - \mathscr{H}^{-1}(\varkappa), \qquad \text{[A.1.18]}$$

where the properties of $\mathscr{H}(p)$ and $\mathscr{V}(p)$ imply that $\mathscr{A}(\varkappa)$ is strictly increasing in $\varkappa$. Note also that $\mathscr{A}(\mathrm{q}) = 1 - 0 = 1$ since $\mathscr{H}(0) = \mathrm{q}$ and $\mathscr{V}(1) = \mathrm{q}$.

Consider first the case where $\underline{\pi} < \overline{\pi}$. The definition [A.1.18] and $\mathscr{V}(\underline{\pi}) = \underline{\varkappa}$ imply:

$$\mathscr{A}(\underline{\varkappa}) = \underline{\pi} - \mathscr{H}^{-1}(\underline{\varkappa}).$$

The definition of $\underline{\pi}$ was such that $\underline{\pi} = 0$ if $\underline{\varkappa} > 0$, and so $\mathscr{A}(\underline{\varkappa}) = -\mathscr{H}^{-1}(\underline{\varkappa})$. Given that $\underline{\varkappa} < \mathrm{q}$, $\mathscr{H}(0) = \mathrm{q}$, and $\mathscr{H}(p)$ is strictly decreasing, it follows that $\mathscr{H}^{-1}(\underline{\varkappa}) > 0$, hence $\mathscr{A}(\underline{\varkappa}) < 0$. When $\underline{\varkappa} = 0$, note that $\mathscr{H}^{-1}(\underline{\varkappa}) = \overline{\pi}$ since $\mathscr{H}(\overline{\pi}) = 0$. This implies that $\mathscr{A}(\underline{\varkappa}) = \underline{\pi} - \overline{\pi} < 0$ because $\underline{\pi} < \overline{\pi}$ in the case under consideration. Therefore, it has been shown that $\mathscr{A}(\underline{\varkappa}) < 0$ and $\mathscr{A}(\mathrm{q}) > 0$, thus $\mathscr{A}(\varkappa)$ being continuous and strictly increasing proves there exists a unique value $\varkappa^* \in (0, \mathrm{q})$ such that $\mathscr{A}(\varkappa^*) = 0$.

Consider the remaining case where $\underline{\pi} \geq \overline{\pi}$. It is necessary that $\underline{\pi} > 0$ in this case since $\overline{\pi} > 0$, and the definition of $\underline{\pi}$ then guarantees that $\mathscr{V}(\underline{\pi}) = 0$, and $\underline{\varkappa} = 0$ because $\underline{\varkappa} = \mathscr{V}(\underline{\pi})$. Therefore, $\mathscr{A}(\underline{\varkappa}) = \mathscr{A}(0) = \underline{\pi} - \overline{\pi} \geq 0$. Since $\mathscr{A}(\varkappa)$ is continuous and strictly increasing and $\mathscr{A}(\mathrm{q}) > 0$ as before, it follows that either $\varkappa^* = 0$ is the only possible solution of $\mathscr{A}(\varkappa) = 0$ for $\varkappa \in [0, \mathrm{q}]$, or there is no solution of the equation. Whether or not a solution exists, set $\varkappa^* = 0$ in this case, noting that $\mathscr{A}(\varkappa^*) \geq 0$.

Depending on which case above prevails, there is either $\mathscr{A}(\varkappa^*) = 0$ or $\mathscr{A}(\varkappa^*) \geq 0$. In both cases, $\varkappa^* < \mathrm{q}$. Define $p^* = \mathscr{H}^{-1}(\varkappa^*)$, and since $\mathscr{H}(p)$ is strictly decreasing and $\mathscr{H}(0) = \mathrm{q}$, it follows that $p^* > 0$. The definition of $\mathscr{H}(p)$ in [A.1.16] also implies that $\mathscr{G}(p^*, \varkappa^*) = 0$. Therefore, using [A.1.12]:

$$\mathrm{q} - \varkappa^* = \frac{p^*}{1 - p^*} u^{-1}(u(\varkappa^*) + \delta). \qquad \text{[A.1.19]}$$

Since $\mathscr{A}(\varkappa^*) \geq 0$, equation [A.1.18] implies $\mathscr{V}^{-1}(\varkappa^*) \geq \mathscr{H}^{-1}(\varkappa^*) = p^*$. As has been shown in [A.1.15], $\mathscr{U}(p, \varkappa)$ is decreasing in p. The definition [A.1.16] implies $\mathscr{U}(\mathscr{V}^{-1}(\varkappa), \varkappa) = 0$, hence it then follows that $\mathscr{U}(p^*, \varkappa^*) \geq 0$. Using [A.1.12]:

$$\mathrm{q} - \varkappa^* \leq \frac{\delta(1 - p^*)}{u'(\varkappa^*)},$$

and combining this with [A.1.19] yields:

$$\frac{p^*}{1 - p^*} u^{-1}(u(\varkappa^*) + \delta) \leq \frac{\delta(1 - p^*)}{u'(\varkappa^*)}.$$

Therefore, the following inequality must hold:

$$u^{-1}(u(\varkappa^*) + \delta) \leq \frac{(1 - p^*)^2}{p^*} \frac{\delta}{u'(C_{\mathrm{w}}^*)}. \qquad \text{[A.1.20]}$$

Now note that since $u(\cdot)$ is a concave function, its inverse $u^{-1}(\cdot)$ is a convex function, so it is bounded below by its tangent at $u(\varkappa^*)$. Together with $\varkappa^* \geq 0$, this leads to:

$$u^{-1}(u(\varkappa^*) + \delta) \geq u^{-1}(u(\varkappa^*)) + \frac{1}{u'(u^{-1}(u(\varkappa^*)))}\delta = \varkappa^* + \frac{\delta}{u'(\varkappa^*)} \geq \frac{\delta}{u'(\varkappa^*)}.$$

By combining this with the earlier inequality in [A.1.20]:

$$\frac{\delta}{u'(\varkappa^*)} \leq \frac{(1 - p^*)^2}{p^*} \frac{\delta}{u'(\varkappa^*)}, \quad \text{and hence } 1 \leq \frac{(1 - p^*)^2}{p^*}.$$

Therefore, the value of p^* must satisfy the quadratic inequality $\mathscr{B}(p^*) \geq 0$ where $\mathscr{B}(p)$ is defined by:

$$\mathscr{B}(p) \equiv (1-p)^2 - p = p^2 - 3p + 1.$$

Since $\mathscr{B}(0) > 0$ and $\mathscr{B}(1) < 0$, the quadratic $\mathscr{B}(p)$ has exactly one root $\overline{p} \in (0,1)$. The product of the roots is positive, so this must be the smallest root, which can then be obtained using the formula:

$$\overline{p} = \frac{3-\sqrt{5}}{2} = 2 - \left(\frac{1+\sqrt{5}}{2}\right) = 2 - \varphi, \qquad \text{[A.1.21]}$$

where $\varphi \equiv (1+\sqrt{5})/2$ is the *Golden ratio* introduced in footnote 11. As $\mathscr{B}(p^*) \geq 0$, it must be the case that $p^* \leq \overline{p}$, and therefore $p^* \leq 2 - \varphi$.

Given the constraint $p < 1/2$, the search for an equilibrium is restricted to the interval $p \in [0, 1/2]$. The non-negativity constraints on consumption are equivalent to $0 \leq C_{\mathrm{w}} \leq \mathrm{q}$. Equation [A.1.2] shows that the value of C_{w} consistent with the binding no-rebellion constraint is strictly decreasing in p. Since the utility function $u(\cdot)$ is strictly increasing and weakly concave, equations [A.1.7] and [A.1.8] imply that any critical point of the objective function C_{p} must be a local maximum. Therefore, these observations show that the general necessary and sufficient conditions for an equilibrium are:

$$\left.\frac{\partial C_{\mathrm{p}}}{\partial p}\right|_{p=p^*} \begin{cases} \leq 0 & \text{if } \ p^* = 0 \ \text{ or } \ C_{\mathrm{w}}^* = \mathrm{q} \\ = 0 & \text{if } \ 0 < p^* < 1/2 \ \text{ and } \ 0 < C_{\mathrm{w}}^* < \mathrm{q} \\ \geq 0 & \text{if } \ p^* = 1/2 \ \text{ or } \ C_{\mathrm{w}}^* = 0 \end{cases}. \qquad \text{[A.1.22]}$$

Consider first the possibility of an equilibrium with $p^* = 0$ or $C_{\mathrm{w}}^* = \mathrm{q}$. Equation [A.1.13] and [A.1.16] imply that $C_{\mathrm{w}}^* = \mathscr{H}(p^*)$, and since $\mathscr{H}(0) = \mathrm{q}$ it follows that any such equilibrium must feature $p^* = 0$ and $C_{\mathrm{w}}^* = \mathrm{q}$. Thus, by using [A.1.12], $\mathscr{U}(p^*, C_{\mathrm{w}}^*) = \delta > 0$. From equation [A.1.13] it follows that $\partial C_{\mathrm{p}}/\partial p \to \infty$ at $p^* = 0$. But the first-order condition [A.1.22] would require $\partial C_{\mathrm{p}}/\partial p \leq 0$ for this type of equilibrium. Therefore, there are no equilibria with either $p^* = 0$ or $C_{\mathrm{w}}^* = \mathrm{q}$.

Now consider the possibility of an equilibrium with $p^* = 1/2$. From the relevant first-order condition in [A.1.22] and [A.1.13], such an equilibrium would need to satisfy $\mathscr{U}(p^*, C_{\mathrm{w}}^*) \geq 0$ and $\mathscr{G}(p^*, C_{\mathrm{w}}^*) = 0$. But it has been shown that $p^* \leq \overline{p}$ for any value of p^* consistent with these conditions, where $\overline{p}$ is defined in [A.1.21]. It can be seen that $\overline{p} < 1/2$, so there are no equilibria with $p^* = 1/2$.

Next, consider the case of an equilibrium with $0 < p^* < 1/2$ and $0 < C_{\mathrm{w}}^* < \mathrm{q}$. Using [A.1.22] and [A.1.13], the required conditions are $\mathscr{G}(p^*, C_{\mathrm{w}}^*) = 0$ and $\mathscr{U}(p^*, C_{\mathrm{w}}^*) = 0$. From [A.1.16], this is seen to be equivalent to $p^* = \mathscr{H}^{-1}(C_{\mathrm{w}}^*)$ and $p^* = \mathscr{V}^{-1}(C_{\mathrm{w}}^*)$, and to $\mathscr{A}(C_{\mathrm{w}}^*) = 0$ using [A.1.18]. In the case where $\underline{\pi} < \overline{\pi}$ such a solution has been shown to exist, and to be unique. There is no solution when $\underline{\pi} \geq \overline{\pi}$. Therefore, an equilibrium of this type exists (and is unique among those of this type) if and only if $\underline{\pi} < \overline{\pi}$.

Finally, consider the case of an equilibrium with $C_{\mathrm{w}}^* = 0$. According to [A.1.13], this must satisfy $\mathscr{G}(p^*, 0) = 0$, and hence $p^* = \mathscr{H}^{-1}(0)$ using [A.1.16]. Using [A.1.22] and [A.1.13], the first-order condition in this case requires $\mathscr{U}(p^*, 0) \geq 0$. The definition in [A.1.16] implies $\mathscr{U}(\mathscr{V}^{-1}(0), 0) = 0$, and since [A.1.15] shows $\mathscr{U}(p, \varkappa)$ is strictly decreasing in p, it follows that $\mathscr{U}(\mathscr{H}^{-1}(0), 0) \geq 0$ if and only if $\mathscr{V}^{-1}(0) \geq \mathscr{H}^{-1}(0)$. This is seen to be equivalent to $\mathscr{A}(\varkappa^*) \geq 0$ using [A.1.18]. The earlier analysis shows this inequality is satisfied if and only if $\underline{\pi} \geq \overline{\pi}$. Hence a unique equilibrium exists in this case too.

Therefore, irrespective of whether $\underline{\pi} < \overline{\pi}$ or $\underline{\pi} \geq \overline{\pi}$ holds, a unique equilibrium exists. In the case $\underline{\pi} < \overline{\pi}$, the equilibrium features $0 < C_{\mathrm{w}}^* < \mathrm{q}$, so all non-negativity constraints are slack. In the case $\underline{\pi} \geq \overline{\pi}$, the equilibrium features $C_{\mathrm{w}}^* = 0$, so the non-negativity constraint is binding for workers. In all cases, $0 < p^* \leq 2 - \varphi < 1/2$. This completes the proof.

A.2 Proof of Proposition 2

Consider an equilibrium $\{p^*, C_{\mathrm{p}}^*(\cdot), C_{\mathrm{w}}^*(\cdot), C_{\mathrm{k}}^*(\cdot)\}$ in which a positive fraction s^* of investment opportunities are taken.

There is a binding no-rebellion constraint with a positive measure of non-incumbents in the rebel army

If not, the power-adjusted fighting effort of all (non-zero measures of) non-incumbents is strictly greater than the power-adjusted fighting effort of incumbents (power adjusted means taking account of the loss of δ when incumbents join the rebel army). Hence it is feasible to reduce consumption of those outside the group in power and distribute it equally among incumbents, which raises incumbent payoffs (the reduction in consumption can be chosen to be small enough so that no-rebellion constraints continue to hold, noting that the number of incumbents willing to fight cannot increase, nor can the power-adjusted fighting effort of those incumbents who are willing to rebel).

Payoff equalization for workers

Suppose there is dispersion in worker consumption $C^*_{\mathrm{w}}(\imath)$ for a positive measure of workers. Let $\zeta(\imath) = C_{\mathrm{k}}(\imath) - C_{\mathrm{w}}(\imath)$ denote the extra consumption allocated to individual $\imath$ if he becomes a capitalist. From [3.6], at the pre-investment stage, those outside the incumbent group have (expected) payoffs $\mathcal{U}_{\mathrm{n}}(\imath) = C_{\mathrm{w}}(\imath) + \alpha \mathbb{E}_{\theta} \max\{\zeta(\imath), 0\}$, while at the post-investment stage, workers have payoffs $\mathcal{U}_{\mathrm{w}}(\imath) = C_{\mathrm{w}}(\imath)$ and capitalists $\mathcal{U}_{\mathrm{k}}(\imath) = C_{\mathrm{w}}(\imath) + \zeta(\imath)$. Now consider a deviation where consumption for every worker is equalized at an amount C_{w} equal to the average of $C^*_{\mathrm{w}}(\imath)$, while $\zeta(\imath)$ and all incumbent payoffs remain unchanged. That reduces payoff inequality among non-incumbents. Since there are more non-incumbents than places in a rebel army (at either stage), the maximum fighting effort from a rebel army including any positive measure of non-incumbents is now strictly lower, while fighting effort from incumbents is unaffected. Since it is known that at least one no-rebellion constraint involving a positive measure of non-incumbents is binding, this argument shows that this binding constraint can be slackened without violating any other constraint. Thus, it must be the case that $C^*_{\mathrm{w}}(\imath) = C^*_{\mathrm{w}}$ in equilibrium.

No capitalists in a rebel army with a binding no-rebellion constraint

Given the incentive constraint and given payoff equalization for workers, capitalists ex post must receive a higher payoff than workers. This means capitalists' power-adjusted fighting effort is less than workers (their power is the same), and since there is never a shortage of workers, capitalists will not feature in a rebel army associated with a binding no-rebellion constraint.

Consumption equalization for capitalists

Since capitalists will not be included in a rebel army with a binding no-rebellion constraint, ex-post inequality among capitalists' payoffs does not increase incentives for rebellion. The question then is how dispersion in the consumption of capitalists affects incentives for investment. An individual $\imath$ with an investment opportunity will invest and become a capitalist if $C_{\mathrm{k}}(\imath) - C_{\mathrm{w}} \geq \theta$. Since $\theta \geq \psi$, a capitalist's consumption cannot be smaller than $C_{\mathrm{w}} + \psi$. Since $\theta \leq \kappa$, a capitalist having consumption larger than $C_{\mathrm{w}} + \kappa$ is costly for the incumbents without yielding any benefit to them. Hence $C_{\mathrm{k}}(\imath) \in [C_{\mathrm{w}} + \psi, C_{\mathrm{w}} + \kappa]$. Using the effort distribution in [3.1], since the utility function $u(\cdot)$ is linear, the probability an individual will choose to take an investment opportunity is $(C_{\mathrm{k}}(\imath) - C_{\mathrm{w}} - \psi)/(\kappa - \psi)$. Averaging that across individuals implies the fraction s of investment opportunities that are taken is given by $(\bar{C}_{\mathrm{k}} - C_{\mathrm{w}} - \psi)/(\kappa - \psi)$, where $\bar{C}_{\mathrm{k}}$ is the average consumption of a capitalist. This shows that given linearity of $u(\cdot)$, dispersion in capitalists' consumption does not affect investment, hence it does not affect any aggregate variable. Thus there is no loss of generality in assuming that all capitalists receive the same consumption C_{k}. Using $s = (C_{\mathrm{k}} - C_{\mathrm{w}} - \psi)/(\kappa - \psi)$ yields the expression in [3.7a].

It follows that all individuals outside the group in power have the same expected payoff:

$$\mathcal{U}_{\mathrm{n}} = C_{\mathrm{w}} + \alpha \mathcal{S}_{\mathrm{i}}(\tilde{\theta}), \quad \text{where } \mathcal{S}_{\mathrm{i}}(\tilde{\theta}) \equiv \mathbb{E}_{\theta} \max\{\tilde{\theta} - \theta, 0\}, \qquad \text{[A.2.1]}$$

where $\mathcal{S}_{\mathrm{i}}(\tilde{\theta})$ is the expected surplus from receiving an investment opportunity. The average incumbent payoff $\bar{\mathcal{U}}_{\mathrm{p}}$ obtained from the budget constraint can then be written as:

$$\bar{\mathcal{U}}_{\mathrm{p}} = \frac{(1-p)\,(\mathrm{q} - \mathcal{U}_{\mathrm{w}}) + \mu(\kappa - \tilde{\theta})s}{p}. \qquad \text{[A.2.2]}$$

The time inconsistency problem

In the equilibrium of the post-investment rebellion subgame, an incumbent's payoff is $\mathcal{U}_{\mathrm{p}}^{\dagger}(K^*)$, where $K^* = \mu\kappa s^*$ is the capital stock if a fraction s^* of investment opportunities are taken up. This payoff is the solution of the problem of maximizing the average incumbent payoff subject to the post-investment no-rebellion constraints. The equilibrium of the subgame features payoff equalization. Now consider a hypothetical problem: maximizing the average incumbent payoff subject to the post-investment stage no-rebellion constraints (with rebels expecting payoffs consistent with the equilibrium) and an additional constraint (but which is not imposed on any subsequent incumbents): a measure μs^* of non-incumbents have to receive consumption $\tilde{\theta} > 0$ above the average of the remaining non-incumbents. The maximized average incumbent payoff $\tilde{\bar{\mathcal{U}}}_{\mathrm{p}}$ in this hypothetical problem must be strictly less than $\mathcal{U}_{\mathrm{p}}^{\dagger}(K)$ (to see this, note that removing this constraint allows resources to be transferred to the incumbents).

Now let $\bar{\mathcal{U}}_{\mathrm{p}}^*$ denote the equilibrium average incumbent payoff. This is the solution to a maximization problem with all of the constraints described in the hypothetical problem above. Note that the capital stock is the same, thus the budget constraint is the same, the extra transfer $\tilde{\theta}$ is required by the incentive constraint for investors, and the no-rebellion constraints in the hypothetical problem are the same as the post-investment stage no-rebellion constraints in the actual problem. In addition, the actual problem also includes the pre-investment stage no-rebellion constraints. It follows that the maximized average incumbent payoff subject to these constraints cannot be more than in the hypothetical problem, hence $\bar{\mathcal{U}}_{\mathrm{p}}^* \leq \tilde{\bar{\mathcal{U}}}_{\mathrm{p}} < \mathcal{U}_{\mathrm{p}}^{\dagger}(K)$.

At least one no-rebellion constraint including only non-incumbents must bind

Since there is no shortage of either non-incumbents at the pre-investment stage or workers at the post-investment stage the following no-rebellion constraints must hold:

$$\mathcal{U}_{\mathrm{p}}' - \mathcal{U}_{\mathrm{n}} \leq \delta \frac{p}{p'}, \quad \text{and} \qquad \text{[A.2.3a]}$$

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{w}} \leq \delta \frac{p}{p^{\dagger}}, \qquad \text{[A.2.3b]}$$

where $p' = p^*$ and $\mathcal{U}_{\mathrm{p}}' = \bar{\mathcal{U}}_{\mathrm{p}}^*$ in equilibrium.

Now suppose for contradiction that neither [A.2.3a] nor [A.2.3b] binds in equilibrium:

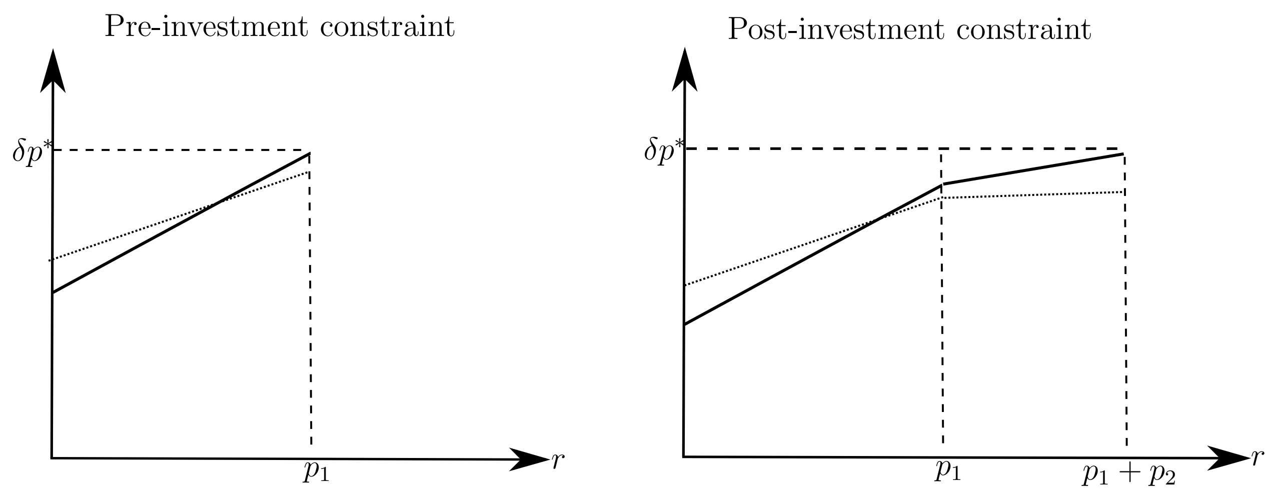
$$\mathcal{U}_{\mathrm{p}}' - \mathcal{U}_{\mathrm{n}} < \delta \frac{p}{p'}, \quad \text{and} \quad \mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{w}} < \delta \frac{p}{p^{\dagger}}.$$

Given whatever consumption distribution is part of the conjectured equilibrium, divide the group in power into three sets. Group 1 would be willing to rebel at either the pre- or post-investment stages (i.e. $\mathcal{U}_{\mathrm{p}}(\imath) \leq \mathcal{U}_{\mathrm{p}}' = \bar{\mathcal{U}}_{\mathrm{p}}^* < \mathcal{U}_{\mathrm{p}}^{\dagger}(K)$). Group 2 would be willing to rebel at the post-investment stage, but not at the pre-investment stage (i.e. $\bar{\mathcal{U}}_{\mathrm{p}}^* = \mathcal{U}_{\mathrm{p}}' < \mathcal{U}_{\mathrm{p}}(\imath) \leq \mathcal{U}_{\mathrm{p}}^{\dagger}(K)$). Group 3 would not rebel at either stage (i.e. $\mathcal{U}_{\mathrm{p}}(\imath) > \mathcal{U}_{\mathrm{p}}^{\dagger}(K)$). In equilibrium, group 1 must contain some positive measure of incumbents, though group 2 and/or group 3 could be empty.

Take the consumption assigned to groups 1 and 2 and redistribute within groups (but not across groups, and group 3 is left untouched) so that all individuals within group 1 have the same payoff and all individuals within group 2 have the same payoff. This redistribution does not directly reduce the incumbents' objective function. Nor does it change the numbers of incumbents who would be willing to rebel at either of the stages. Furthermore, it cannot increase the effective fighting strength that can be obtained from the strongest rebel army conditional on including a given measure of incumbents (effective fighting strength is fighting effort with δ added when an incumbent joins a rebel army).

Let p_1 denote the measure of incumbents in group 1, and p_2 the measure in group 2. Following the redistribution, the incumbent payoff distribution has two point masses at $\mathcal{U}_{\mathrm{p},1}$ and $\mathcal{U}_{\mathrm{p},2}$, and potentially a general distribution above $\mathcal{U}_{\mathrm{p}}^{\dagger}(K)$. Figure 4 depicts total effective fighting strength as a function of the measure r of incumbents included in a rebel army, for both pre- and post-investment stages. These functions

Figure 4: *Total effective fighting strength*



are piecewise linear, bounded above by δp^*, and given the supposition, both functions are strictly below δp^* at $r = 0$. The first function is defined only for $r \in [0, p_1]$, and the second is defined for $r \in [0, p_1 + p_2]$.

Consider a further deviation. Reduce consumption of all non-incumbents by an amount ϵ, distributing the resources equally among incumbents in group 1. The ϵ is chosen so that both no-rebellion constraints remain slack at $r = 0$. Hence the group-1 incumbents obtain an increment $\epsilon(1 - p^*)/p_1$ to consumption and utility. The functions in Figure 4 remain piecewise linear following this. At $r = p_1$, the total effective fighting strength of group-1 incumbents in the rebel army is reduced by $(1 - p^*)\epsilon$, while the effect on workers' total fighting effort is $(p^* - p_1)\epsilon$. The latter is smaller, hence the net effect is negative. Since the function corresponding to the pre-investment constraint is linear, that no-rebellion constraint must be slack. For the function corresponding to the post-investment no-rebellion constraint, note that the slope between p_1 and $p_1 + p_2$ is also reduced because the gap between $\mathcal{U}_{\mathrm{w}}$ and $\mathcal{U}_{\mathrm{p},2}$ is increased (the slope is equal to $\delta + \mathcal{U}_{\mathrm{w}} - \mathcal{U}_{\mathrm{p},2}$). Hence, the post-investment constraint is now also slack. Since the average incumbent payoff has been increased by the deviation, the initial supposition must be false.

Power sharing must be larger than it would be in equilibrium following a rebellion at the post-investment stage

Suppose not for contradiction, that is, $p^* \leq p^\dagger$. It is known that one of the constraints [A.2.3a] and [A.2.3b] must be binding. Consider the case where it is [A.2.3a]. This means that in equilibrium ($p' = p^*$):

$$\bar{\mathcal{U}}_{\mathrm{p}}^* - \mathcal{U}_{\mathrm{n}}^* = \delta.$$

Using this equation and [A.2.1], it follows that:

$$\mathcal{U}_{\mathrm{p}}^\dagger(K) - \mathcal{U}_{\mathrm{w}}^* = (\mathcal{U}_{\mathrm{p}}^\dagger(K) - \bar{\mathcal{U}}_{\mathrm{p}}^*) + \mathcal{S}_{\mathrm{i}}(\tilde{\theta}) + \delta > \delta \geq \delta \frac{p^*}{p^\dagger},$$

where the first inequality uses the time inconsistency result and the second uses the conjecture $p^* \leq p^\dagger$. It follows that the no-rebellion constraint [A.2.3b] would be violated in this case.

Therefore, if $p^* \leq p^\dagger$ then it must be the case that [A.2.3b] binds:

$$\mathcal{U}_{\mathrm{p}}^\dagger(K) - \mathcal{U}_{\mathrm{w}}^* = \delta \frac{p^*}{p^\dagger}.$$

This represents a rebellion where $p^\dagger$ workers each put in $\delta p^*/p^\dagger$ units of fighting effort, bringing the total fighting effort exactly to δp^*. However, it is known that $\bar{\mathcal{U}}_{\mathrm{p}}^* < \mathcal{U}_{\mathrm{p}}^\dagger(K)$. Obviously, since $\bar{\mathcal{U}}_{\mathrm{p}}^*$ is an average, there must be a positive measure of incumbents receiving this average payoff or a lower one. These incumbents have a strict preference to rebel at the post-investment stage. If they joined a rebellion, their power-adjusted fighting strength would be δ plus some strictly positive number. Hence substituting a positive measure of these incumbents into the rebel army replaces workers exerting fighting effort $\delta p^*/p^\dagger \leq \delta$ with incumbents with power-adjusted fighting strength strictly larger than δ. Since the constraint holds exactly with only workers, it is now violated once the rebel army includes a positive measure of incumbents. Therefore, contradictions are obtained in all cases supposing $p^* \leq p^\dagger$, hence it is shown that the equilibrium must feature $p^* > p^\dagger$.

The pre-investment stage constraint for a rebel army with only non-incumbents cannot bind on its own

Consider the possibility of an equilibrium with $s > 0$ in which the pre-investment stage constraint [A.2.3a] is the only effective binding no-rebellion constraint. With $p' = p^*$ and $\mathcal{U}'_{\mathrm{p}} = \mathcal{U}^*_{\mathrm{p}}$, the binding constraint [A.2.3a] implies

$$\mathcal{U}^*_{\mathrm{p}} - \mathcal{U}_{\mathrm{n}} = \delta \frac{p}{p^*}.$$

Combining this with the expressions for the non-incumbent expected payoff $\mathcal{U}_{\mathrm{n}}$ from [A.2.1] (using $\mathcal{U}_{\mathrm{w}} = C_{\mathrm{w}}$) yields:

$$\mathcal{U}_{\mathrm{w}} = \mathcal{U}^*_{\mathrm{p}} - \delta \frac{p}{p^*} - \alpha \mathcal{S}_{\mathrm{i}}(\tilde{\theta}) = \mathcal{U}^*_{\mathrm{p}} - \delta \frac{p}{p^*} - \frac{\mu}{1-p} \mathcal{S}_{\mathrm{i}}(\tilde{\theta}), \tag{A.2.4}$$

where the formula for the probability α of receiving an investment opportunity is taken from [3.6]. Using the expression for $\bar{\mathcal{U}}_{\mathrm{p}}$ from [A.2.2] and the formula for $\mathcal{U}_{\mathrm{w}}$ yields:

$$\bar{\mathcal{U}}_{\mathrm{p}} = \frac{(1-p)\left(\mathrm{q} + \delta \frac{p}{p^*} - \mathcal{U}^*_{\mathrm{p}}\right) + \mu(\kappa - \tilde{\theta})s + \mu \mathcal{S}_{\mathrm{i}}(\tilde{\theta})}{p}. \tag{A.2.5}$$

The first-order condition for maximizing $\bar{\mathcal{U}}_{\mathrm{p}}$ with respect to p (taking p^* and $\mathcal{U}^*_{\mathrm{p}}$ as given) is:

$$\frac{1}{p}\left((1-p)\left(\mathrm{q} + \delta \frac{p}{p^*} - \mathcal{U}^*_{\mathrm{p}}\right) + \mu(\kappa - \tilde{\theta})s + \mu \mathcal{S}_{\mathrm{i}}(\tilde{\theta})\right) = (1-p)\frac{\delta}{p^*} - \left(\mathrm{q} + \delta \frac{p}{p^*} - \mathcal{U}^*_{\mathrm{p}}\right),$$

and after imposing the equilibrium conditions $p = p^*$ and $\bar{\mathcal{U}}_{\mathrm{p}} = \mathcal{U}^*_{\mathrm{p}}$ and making use of [A.2.5]:

$$\mathcal{U}^*_{\mathrm{p}} = \delta \frac{(1-p^*)}{p^*} - \left(\mathrm{q} + \delta - \mathcal{U}^*_{\mathrm{p}}\right).$$

Solving this equation for p^* yields:

$$p^* = \frac{\delta}{\mathrm{q} + 2\delta} = p^\dagger,$$

given the expression for $p^\dagger$ in [3.5]. This contradicts the requirement that $p^* > p^\dagger$, so this constraint cannot bind on its own.

A binding pre-investment stage constraint for non-incumbents requires payoff equalization for the incumbents and implies that other pre-investment stage constraints may be ignored

If constraint [A.2.3a] is binding in equilibrium then:

$$\bar{\mathcal{U}}^*_{\mathrm{p}} - \mathcal{U}^*_{\mathrm{n}} = \delta.$$

In this rebel army, a measure p^* of non-incumbents each exert δ fighting effort. Now suppose there is payoff inequality among a positive measure of incumbents in this case. Since $\bar{\mathcal{U}}^*_{\mathrm{p}}$ is the average payoff, this means there is a positive measure of incumbents with payoffs strictly lower than $\bar{\mathcal{U}}_{\mathrm{p}}^{*}$. As $\mathcal{U}_{\mathrm{p}}^{\prime}=\bar{\mathcal{U}}_{\mathrm{p}}^{*}$ holds, it follows that this positive measure of incumbents would like to join a rebellion at the pre-investment stage and exert a strictly positive amount of fighting effort. Hence they would contribute power-adjusted fighting strength strictly larger than δ. This positive measure in combination with non-incumbents exerting effort δ would lead to a total power-adjusted fighting strength of the rebel army greater than δp^{*}, and hence a successful rebellion. So in the case under consideration, there must be complete payoff equalization among the incumbents.

Following the argument in Proposition 1, the set of equilibria imposing only the non-incumbents' pre-investment stage no-rebellion constraint is the same as the set of equilibria imposing all combinations of pre-investment stage no-rebellion constraints.

The post-investment stage constraint involving only workers cannot bind on its own

The expression for the average incumbent payoff from [A.2.2] can be rearranged as follows:

$$\bar{\mathcal{U}}_{\mathrm{p}}=\frac{(1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)\right)+\mu(\kappa-\tilde{\theta})s}{p}-\left(\frac{1-p}{p}\right)\left(\mathcal{U}_{\mathrm{w}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)+\delta\frac{p}{p^{\dagger}}\right),$$

from which an expression for $\mathcal{U}_{\mathrm{p}}^{\dagger}(K)-\bar{\mathcal{U}}_{\mathrm{p}}$ can be obtained:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K)-\bar{\mathcal{U}}_{\mathrm{p}}=\frac{1}{p}\left(\mathcal{U}_{\mathrm{p}}^{\dagger}(K)-\left((1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}\right)+\mu(\kappa-\tilde{\theta})s\right)\right)+\left(\frac{1-p}{p}\right)\left(\mathcal{U}_{\mathrm{w}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)+\delta\frac{p}{p^{\dagger}}\right). \quad \text{[A.2.6]}$$

Substituting for $\mathcal{U}_{\mathrm{p}}^{\dagger}(K)$ from [3.5] and rearranging the first term in brackets yields:

$$\begin{aligned}
\mathcal{U}_{\mathrm{p}}^{\dagger}(K)-\left((1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}\right)+\mu(\kappa-\tilde{\theta})s\right)&=\frac{(\mathrm{q}+\delta)^{2}}{\mathrm{q}+2\delta}+K-(1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}\right)-\mu\kappa s+\mu\tilde{\theta}s\\
&=(\mathrm{q}+\delta)(1-p^{\dagger})+\mu\kappa s-\mathrm{q}(1-p)-\frac{\delta}{p^{\dagger}}p+\frac{\delta}{p^{\dagger}}p^{2}-\mu\kappa s+\mu\tilde{\theta}s\\
&=\frac{\delta}{p^{\dagger}}\left(p^{2}-p+\frac{\mathrm{q}p^{\dagger}}{\delta}p+\frac{p^{\dagger}}{\delta}\left((\mathrm{q}+\delta)(1-p^{\dagger})-\mathrm{q}\right)\right)+\mu\tilde{\theta}s\\
&=\frac{\delta}{p^{\dagger}}\left(p^{2}-\left(1-\frac{\mathrm{q}}{\mathrm{q}+2\delta}\right)p+\frac{p^{\dagger}}{\delta}\left(\frac{(\mathrm{q}+\delta)^{2}-\mathrm{q}(\mathrm{q}+2\delta)}{\mathrm{q}+2\delta}\right)\right)+\mu\tilde{\theta}s\\
&=\frac{\delta}{p^{\dagger}}\left(p^{2}-2\left(\frac{\delta}{\mathrm{q}+2\delta}\right)p+p^{\dagger}\left(\frac{\delta}{\mathrm{q}+2\delta}\right)\right)+\mu\tilde{\theta}s\\
&=\frac{\delta}{p^{\dagger}}\left(p^{2}-2p^{\dagger}p+p^{\dagger 2}\right)+\mu\tilde{\theta}s=\frac{\delta}{p^{\dagger}}\left(p-p^{\dagger}\right)^{2}+\mu\tilde{\theta}s.
\end{aligned} \quad \text{[A.2.7]}$$

Now consider the possibility of an equilibrium featuring $s>0$ with workers' post-investment constraint [A.2.3b] as the only effective binding no-rebellion constraint. When the constraint in [A.2.3b] binds:

$$\mathcal{U}_{\mathrm{w}}=\mathcal{U}_{\mathrm{p}}^{\dagger}(K)-\delta\frac{p}{p^{\dagger}}, \quad \text{[A.2.8]}$$

which is substituted into the expression for the average incumbent payoff [A.2.2] to obtain:

$$\bar{\mathcal{U}}_{\mathrm{p}}=\frac{(1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)\right)+\mu(\kappa-\tilde{\theta})s}{p}. \quad \text{[A.2.9]}$$

Taking the derivative of $\bar{\mathcal{U}}_{\mathrm{p}}$ with respect to p:

$$\frac{\partial\bar{\mathcal{U}}_{\mathrm{p}}}{\partial p}=\frac{1}{p}\left((1-p)\frac{\delta}{p^{\dagger}}-\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)\right)\right)-\frac{1}{p^{2}}\left((1-p)\left(\mathrm{q}+\delta\frac{p}{p^{\dagger}}-\mathcal{U}_{\mathrm{p}}^{\dagger}(K)\right)+\mu(\kappa-\tilde{\theta})s\right),$$

and by using the expressions for $\bar{\mathcal{U}}_{\mathrm{p}}$ and $p^{\dagger}$ from [A.2.9] and [3.5] respectively:

$$\frac{\partial \bar{\mathcal{U}}_{\mathrm{p}}}{\partial p} = \frac{1}{p}\left(\frac{\delta}{p^{\dagger}} - \mathrm{q} - 2\delta\frac{p}{p^{\dagger}} + (\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}})\right) = \frac{1}{p}\left(2\frac{\delta}{p^{\dagger}}(p^{\dagger} - p) + (\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}})\right). \quad \text{[A.2.10]}$$

Exploiting the fact that [A.2.8] holds in this case, equations [A.2.6] and [A.2.7] imply that:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}} = \frac{1}{p}\left(\frac{\delta}{p^{\dagger}}(p - p^{\dagger})^2 + \mu\tilde{\theta}s\right), \quad \text{[A.2.11]}$$

and by substituting this into [A.2.10]:

$$\frac{\partial \bar{\mathcal{U}}_{\mathrm{p}}}{\partial p} = \frac{1}{p^2}\left(2\frac{\delta}{p^{\dagger}}(p^{\dagger} - p)p + \frac{\delta}{p^{\dagger}}(p - p^{\dagger})^2 + \mu\tilde{\theta}s\right) = \frac{\delta}{p^2 p^{\dagger}}\left(p^{\dagger}\left(p^{\dagger} + \frac{\mu\tilde{\theta}s}{\delta}\right) - p^2\right). \quad \text{[A.2.12]}$$

Now make the following definition of the function $\pi(s)$:

$$\pi(s) \equiv p^{\dagger} + \frac{\mu\tilde{\theta}s}{\delta}. \quad \text{[A.2.13]}$$

The restrictions in [3.3] imply that $\pi(s) < 1/2$ for all $s \in [0, 1]$. Using the definition of $\pi(s)$ in [A.2.13] and equation [A.2.12], the derivative of $\bar{\mathcal{U}}_{\mathrm{p}}$ can be written as:

$$\frac{\partial \bar{\mathcal{U}}_{\mathrm{p}}}{\partial p} = \frac{\delta}{p^2 p^{\dagger}}\left(p^{\dagger}\pi(s) - p^2\right). \quad \text{[A.2.14]}$$

It follows that $\bar{\mathcal{U}}_{\mathrm{p}}$ is strictly increasing in p for $p < \sqrt{p^{\dagger}}\sqrt{\pi(s)}$, and strictly decreasing for $p > \sqrt{p^{\dagger}}\sqrt{\pi(s)}$. The first-order condition for maximizing the expression in [A.2.9] for $\bar{\mathcal{U}}_{\mathrm{p}}$ incorporating the binding constraint is therefore $p^* = \sqrt{p^{\dagger}}\sqrt{\pi(s)}$. It can be seen from [A.2.13] that $p^{\dagger} < \pi(s)$ for any $s > 0$, so it follows that $p^{\dagger} < p^* < \pi(s)$.

In the case under consideration where [A.2.3b] is binding, equation [A.2.11] implies that the inequality

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}} \leq \delta\left(\frac{p - p^{\dagger}}{p^{\dagger}}\right) \quad \text{[A.2.15]}$$

is equivalent to the following:

$$\frac{\delta}{p^{\dagger}}(p - p^{\dagger})^2 + \mu s\tilde{\theta} \leq \frac{\delta}{p^{\dagger}}(p - p^{\dagger})p.$$

Simplification of the terms appearing in the inequality above shows that it is in turn equivalent to $p \geq \pi(s)$, with $\pi(s)$ as defined in [A.2.13]. But the first-order condition $p^* = \sqrt{p^{\dagger}}\sqrt{\pi(s)}$ implies $p^* < \pi(s)$ since $\pi(s) > p^{\dagger}$ for $s > 0$. Therefore:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}}^* > \delta\left(\frac{p^* - p^{\dagger}}{p^{\dagger}}\right).$$

Given the constraint [A.2.3b] that is assumed to bind, a rebellion comprising $p^{\dagger}$ workers would have each exert fighting effort $\delta p^*/p^{\dagger}$, which adds up to the total fighting strength δp^* of the incumbent army.

Now note that whatever the distribution of incumbent payoffs, since $\bar{\mathcal{U}}_{\mathrm{p}}^*$ is the average, there must be a strictly positive measure of incumbents receiving a payoff of $\bar{\mathcal{U}}_{\mathrm{p}}^*$ or less. Combined with the inequality derived above, it follows that this positive measure of incumbents is willing to rebel, and the inequality reveals that each of that group would be willing to put in an amount of fighting effort strictly larger than $\delta p^*/p^{\dagger} - \delta$. Since each has power δ, all would contribute power-adjusted fighting strength strictly greater than $\delta p^*/p^{\dagger}$. Now taking some of these incumbents and combining them with workers leads to a rebel army with total power-adjusted fighting strength more than δp^*, so the rebellion would succeed. Hence it is not possible that [A.2.3b] is the only binding constraint.

The pre-investment stage constraint with only non-incumbents cannot bind in conjunction with a post-investment constraint involving a positive number of incumbents

It has been shown that if [A.2.3a] binds in equilibrium then all incumbents must receive the same payoff. Let $\mathcal{U}_{\mathrm{p}}^{*}$ denote this common payoff. Given that all workers receive the same payoff and the no-rebellion constraint [A.2.3b] must hold, and since $p^{*} > p^{\dagger}$, if a post-investment stage no-rebellion constraint is binding with a positive measure of incumbents, it must be that the no-rebellion constraint including only incumbents is binding (the argument is that total effective fighting strength is linear in the composition of the rebel army):

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{p}}^{*} = \delta\left(\frac{p^{*} - p^{\dagger}}{p^{\dagger}}\right).$$

Using equation [3.6] and the binding constraint [A.2.3a], it follows that $\mathcal{U}_{\mathrm{w}}^{*} = \mathcal{U}_{\mathrm{p}}^{*} - \alpha\mathcal{S}_{\mathrm{i}}(\tilde{\theta}) - \delta$. The equation above implies $\mathcal{U}_{\mathrm{p}}^{\dagger}(K) = \mathcal{U}_{\mathrm{p}}^{*} + \delta p^{*}/p^{\dagger} - \delta$, and putting these two results together yields:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{w}}^{*} = \delta\frac{p^{*}}{p^{\dagger}} + \alpha\mathcal{S}_{\mathrm{i}}(\tilde{\theta}),$$

but since $\alpha\mathcal{S}_{\mathrm{i}}(\tilde{\theta}) > 0$, this violates [A.2.3b]. Therefore, this combination of binding constraints is inconsistent with equilibrium.

The pre-investment stage constraint with only non-incumbents cannot bind in conjunction with the post-investment stage constraint with only workers

Since the pre-investment stage no-rebellion constraint for non-incumbents is binding, there must be complete payoff equalization among incumbents. Furthermore, when this constraint binds, an expression for $\bar{\mathcal{U}}_{\mathrm{p}}$ is given in [A.2.5]. After imposing the equilibrium conditions $p = p^{*}$ and $\mathcal{U}_{\mathrm{p}}^{*} = \bar{\mathcal{U}}_{\mathrm{p}}$ in that equation:

$$\mathcal{U}_{\mathrm{p}}^{*} = (\mathrm{q} + \delta)(1 - p^{*}) + \mu(\kappa - \tilde{\theta})s + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}).$$

It has already been shown that when [A.2.3a] binds, equation [A.2.4] for $\mathcal{U}_{\mathrm{w}}$ follows. By substituting the equation above into [A.2.4]:

$$\mathcal{U}_{\mathrm{w}}^{*} = \mathrm{q}(1 - p^{*}) - \delta p^{*} + \mu(\kappa - \tilde{\theta})s - \mu\frac{p^{*}}{1 - p^{*}}\mathcal{S}_{\mathrm{i}}(\tilde{\theta}).$$

This can be substituted into [A.2.3b] to obtain a condition for p^{*} to satisfy the post-investment no-rebellion constraint for workers:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathrm{q}(1 - p^{*}) + \delta p^{*} - \mu(\kappa - \tilde{\theta})s + \mu\frac{p^{*}}{1 - p^{*}}\mathcal{S}_{\mathrm{i}}(\tilde{\theta}) \leq \delta\frac{p^{*}}{p^{\dagger}}.$$

Using the expressions for $\mathcal{U}_{\mathrm{p}}^{\dagger}(K)$ and $p^{\dagger}$ from [3.5], and $K = \mu\kappa s$, it can be seen that $\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathrm{q} = \delta p^{\dagger} + \mu\kappa s$. By substituting this into the inequality above it follows that it is equivalent to

$$\delta p^{\dagger} + (\mathrm{q} + \delta)p^{*} + \mu\tilde{\theta}s + \mu\frac{p^{*}}{1 - p^{*}}\mathcal{S}_{\mathrm{i}}(\tilde{\theta}) \leq \frac{\delta}{p^{\dagger}}p^{*}.$$

Noting that [3.5] implies $\delta/p^{\dagger} = \mathrm{q} + 2\delta$, the inequality above can be rearranged to yield

$$\left(1 - \frac{\mu}{1 - p^{*}}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\theta})}{\delta}\right)p^{*} \geq \pi(s), \qquad \text{[A.2.16a]}$$

where $\pi(s)$ is defined in [A.2.13].

Now start from the case where the post-investment no-rebellion constraint for workers [A.2.3b] is binding at $p = p^*$. The condition needed for an equilibrium ($p' = p^*$ and $\mathcal{U}'_\mathrm{p} = \mathcal{U}^*_\mathrm{p}$) of this type to satisfy the pre-investment stage no-rebellion constraint [A.2.3a] is

$$\mathcal{U}^*_\mathrm{p} - \mathcal{U}^*_\mathrm{n} \leq \delta.$$

The expected payoff $\mathcal{U}_\mathrm{n}$ for non-incumbents is given in equation [A.2.1]. Using the formula for α from [3.6], it follows that $\mathcal{U}^*_\mathrm{n} = \mathcal{U}^*_\mathrm{w} + (\mu/(1-p^*))\mathcal{S}_\mathrm{i}(\tilde{\theta})$. Substituting this into the inequality above shows that [A.2.3a] is satisfied if and only if

$$\mathcal{U}^*_\mathrm{p} - \mathcal{U}^*_\mathrm{w} - \frac{\mu}{1-p^*}\mathcal{S}_\mathrm{i}(\tilde{\theta}) \leq \delta.$$

A binding post-investment constraint [A.2.3b] for workers implies that $\mathcal{U}^*_\mathrm{w} = \mathcal{U}^\dagger_\mathrm{p}(K) - \delta p^*/p^\dagger$, which can be substituted into the condition above to obtain:

$$\delta\frac{p^*}{p^\dagger} - \delta - \frac{\mu}{1-p^*}\mathcal{S}_\mathrm{i}(\tilde{\theta}) \leq \mathcal{U}^\dagger_\mathrm{p}(K) - \mathcal{U}^*_\mathrm{p}.$$

Equation [A.2.8] (the binding version of [A.2.3b]) has already been shown to imply the expression for $\mathcal{U}^\dagger_\mathrm{p}(K) - \bar{\mathcal{U}}_\mathrm{p}$ in [A.2.11]. Since $\bar{\mathcal{U}}^*_\mathrm{p} = \mathcal{U}^*_\mathrm{p}$, this can be used to deduce that the inequality above is equivalent to:

$$\frac{\delta}{p^\dagger}(p^* - p^\dagger)p^* - \mu\frac{p^*}{1-p^*}\mathcal{S}_\mathrm{i}(\tilde{\theta}) \leq \frac{\delta}{p^\dagger}(p^* - p^\dagger)^2 + \mu\tilde{\theta}s.$$

After some rearrangement, this condition can be written as

$$p^* - p^\dagger - \mu\frac{p^*}{1-p^*}\frac{\mathcal{S}_\mathrm{i}(\tilde{\theta})}{\delta} \leq \frac{\mu\tilde{\theta}s}{\delta},$$

which can be stated in terms of the function $\pi(s)$ from [A.2.13]:

$$\left(1 - \frac{\mu}{1-p^*}\frac{\mathcal{S}_\mathrm{i}(\tilde{\theta})}{\delta}\right)p^* \leq \pi(s). \qquad \text{[A.2.16b]}$$

Note that if both constraints [A.2.3a] and [A.2.3b] are binding in equilibrium then both [A.2.16a] and [A.2.16b] must hold, hence:

$$\left(1 - \frac{\mu}{1-p^*}\frac{\mathcal{S}_\mathrm{i}(\tilde{\theta})}{\delta}\right)p^* = \pi(s). \qquad \text{[A.2.17]}$$

The expected surplus to those receiving an investment opportunity is $\mathcal{S}_\mathrm{i}(\tilde{\theta})$, which is defined in equation [A.2.1]. The probability distribution of the effort cost θ in [3.1] has density function $1/(\kappa - \psi)$ on support $[\psi, \kappa]$, so an explicit expression for the surplus is:

$$\mathcal{S}_\mathrm{i}(\tilde{\theta}) = \int_{\theta=\psi}^{\tilde{\theta}} \frac{\tilde{\theta} - \theta}{\kappa - \psi}\mathrm{d}\theta = \frac{1}{2}\frac{(\tilde{\theta} - \psi)^2}{\kappa - \psi}. \qquad \text{[A.2.18]}$$

The parameter restrictions in [3.3] require $\kappa < \delta$. Together with $\psi \leq \tilde{\theta} \leq \kappa$ and $0 < \psi < \kappa$, this implies

$$\mathcal{S}_\mathrm{i}(\tilde{\theta}) \leq \frac{1}{2}(\kappa - \psi) < \frac{\kappa}{2} < \frac{\delta}{2}. \qquad \text{[A.2.19]}$$

Now define the function $\mathscr{M}(p)$ as follows and calculate its derivative:

$$\mathscr{M}(p) \equiv \left(1 - \frac{\upmu}{1-p}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}\right)p, \quad \text{and} \quad \mathscr{M}'(p) = 1 - \frac{1}{1-p}\frac{\mu}{1-p}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}. \qquad \text{[A.2.20]}$$

Given the parameter restrictions in [3.3], the formula $\alpha = \upmu/(1-p)$ from [3.6] always returns a well-defined probability for $p \leq 1/2$. Given that $p \leq 1/2$ implies $1/(1-p) \leq 2$ and [A.2.19] implies $\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})/\updelta < 1/2$, it follows from [A.2.20] that $\mathscr{M}'(p) > 0$ for all $p \in [0, 1/2]$. Noting that $\mathscr{M}(0) = 0$ and $\pi(s) > 0$ according to [A.2.13], the equation $\mathscr{M}(p^*) = \pi(s)$ has at most one solution p^* satisfying $0 < p^* < 1/2$. When such a solution exists, it follows from [A.2.20] that $p^* > \pi(s)$ because $\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})/\updelta < 1/2$ and $\alpha = \upmu/(1-p^*) \leq 1$. Observe that the equation $\mathscr{M}(p^*) = \pi(s)$ is equivalent to the condition in [A.2.17] for both constraints to be binding.

In the case where no solution $p^* \in (0, 1/2)$ exists, set $p^* = 1/2$, and as $\mathscr{M}(0) = 0$ and $\mathscr{M}'(p) > 0$ it must be the case that $\mathscr{M}(p^*) = \mathscr{M}(1/2) < \pi(s)$. Since it has been shown that $\pi(s) < 1/2$ for the function $\pi(s)$ defined in [A.2.13], it follows that $p^* > \pi(s)$. Note that this case is consistent with [A.2.16b], but not [A.2.16a], which means that it could only occur in conjunction with the post-investment no-rebellion constraint for workers being binding, while the pre-investment no-rebellion constraint is slack.

Now consider a deviation from either of the conjectured equilibria described above, namely power sharing satisfying $0 < p^* < 1/2$ and both no-rebellion constraints binding, or $p^* = 1/2$ with the pre-investment no-rebellion constraint slack. Workers' post-investment no-rebellion constraint binds in both cases. The deviation involves changing power sharing p, while consumption of workers is adjusted so that the post-investment no-rebellion constraint for workers continues to bind, in which case the payoff of a worker is given in equation [A.2.8]. The condition for the new choice of p to be consistent with the pre-investment stage constraint (with $p' = p^*$ and $\mathcal{U}'_{\mathrm{p}} = \mathcal{U}^*_{\mathrm{p}}$ taken as given at their conjectured equilibrium values) is:

$$\mathcal{U}^*_{\mathrm{p}} - \mathcal{U}_{\mathrm{n}} \leq \updelta \frac{p}{p^*}, \qquad \text{[A.2.21]}$$

where the expected non-incumbent payoff is obtained using equations [A.2.1] and [A.2.8]:

$$\mathcal{U}_{\mathrm{n}} = \mathcal{U}_{\mathrm{w}} + \frac{\upmu}{1-p}\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta}) = \mathcal{U}^{\dagger}_{\mathrm{p}}(K) - \updelta\frac{p}{p^{\dagger}} + \frac{\upmu}{1-p}\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta}).$$

Substituting this expression into [A.2.21] shows that the feasibility of the deviation requires:

$$\mathcal{U}^*_{\mathrm{p}} - \mathcal{U}^{\dagger}_{\mathrm{p}}(K) + \updelta\frac{p}{p^{\dagger}} - \frac{\upmu}{1-p}\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta}) \leq \updelta\frac{p}{p^*},$$

which can be rearranged as follows:

$$\left(\frac{1}{p^{\dagger}} - \frac{1}{p^*}\right)p - \frac{\upmu}{1-p}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta} \leq \frac{\mathcal{U}^{\dagger}_{\mathrm{p}}(K) - \mathcal{U}^*_{\mathrm{p}}}{\updelta}. \qquad \text{[A.2.22]}$$

This inequality is known to be satisfied at $p = p^*$ because p^* satisfies the pre-investment no-rebellion constraint.

Now define the function $\mathscr{K}(p)$, and note the following expression for its derivative:

$$\mathscr{K}(p) \equiv \left(\frac{1}{p^{\dagger}} - \frac{1}{p^*}\right)p - \frac{\upmu}{1-p}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}, \quad \text{and} \quad \mathscr{K}'(p) = \frac{1}{p^{\dagger}} - \frac{1}{p^*} - \frac{\upmu}{(1-p)^2}\frac{\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}. \qquad \text{[A.2.23]}$$

Evaluating the derivative at $p = p^*$:

$$\mathscr{K}'(p^*) = \frac{1}{p^{\dagger}p^*}\left(p^* - p^{\dagger} - \frac{\upmu\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}\frac{p^{\dagger}}{1-p^*}\frac{p^*}{1-p^*}\right) > \frac{1}{p^{\dagger}p^*}\left(\pi(s) - p^{\dagger} - \frac{\upmu\mathcal{S}_{\mathrm{i}}(\tilde{\uptheta})}{\updelta}\frac{p^{\dagger}}{1-p^*}\frac{p^*}{1-p^*}\right),$$

where the inequality above follows from $p^* > \pi(s)$, which is true in all cases under consideration. Since [A.2.13] implies $\pi(s) - p^\dagger = \mu\tilde{\theta}s/\delta$, the inequality above implies:

$$\mathscr{K}'(p^*) > \frac{\mu}{\delta}\frac{1}{p^\dagger p^*}\left(\tilde{\theta}s - \frac{p^\dagger}{1-p^*}\frac{p^*}{1-p^*}\mathcal{S}_\mathrm{i}(\tilde{\theta})\right). \qquad \text{[A.2.24]}$$

Note that $s = \mathbb{P}_\theta[\theta \leq \tilde{\theta}]$, and use the definition of $\mathcal{S}_\mathrm{i}(\tilde{\theta})$ in [A.2.1] to deduce

$$\tilde{\theta}s - \mathcal{S}_\mathrm{i}(\tilde{\theta}) = \mathbb{E}_\theta\tilde{\theta}\mathbb{1}[\theta \leq \tilde{\theta}] - \mathbb{E}_\theta \max\{\tilde{\theta} - \theta, 0\} = \mathbb{E}_\theta\theta\mathbb{1}[\theta \leq \tilde{\theta}] \geq 0,$$

and hence $\mathcal{S}_\mathrm{i}(\tilde{\theta}) \leq \tilde{\theta}s$. Substituting this into [A.2.24] demonstrates that

$$\mathscr{K}'(p^*) > \frac{\mu}{\delta}\frac{\tilde{\theta}s}{p^\dagger p^*}\left(1 - \frac{p^\dagger}{1-p^*}\frac{p^*}{1-p^*}\right).$$

Since $p^\dagger < p^* \leq 1/2$, it follows that $p^\dagger/(1-p^*) < 1$ and $p^*/(1-p^*) \leq 1$, and thus $\mathscr{K}'(p^*) > 0$.

The pre-investment no-rebellion constraint [A.2.21] is equivalent to [A.2.22], which by comparison with the definition of $\mathscr{K}(p)$ in [A.2.23] is in turn equivalent to:

$$\mathscr{K}(p) \leq \frac{\mathcal{U}_\mathrm{p}^\dagger(K) - \mathcal{U}_\mathrm{p}^*}{\delta}.$$

It is known that $\mathscr{K}(p)$ is increasing in p and the above inequality is satisfied at $p = p^*$. Observing that the right-hand side is unaffected by movements in p away from the conjectured equilibrium p^*, it then follows that it is possible to reduce p below p^* and still satisfy the pre-investment no-rebellion constraint.

Now consider the incumbents' post-investment stage no-rebellion constraint. Since incumbent payoffs must be equalized, the inequality in [A.2.15] is equivalent to that no-rebellion constraint. Furthermore, it has already been shown that when the post-investment stage no-rebellion constraint binds for workers, [A.2.15] is equivalent to $p \geq \pi(s)$.

Therefore, starting from the conjectured equilibrium at $p = p^* > \pi(s)$, it is feasible to reduce p by some positive amount and continue to satisfy all no-rebellion constraints. Moreover, it has been shown that when the workers' binding post-investment constraint is used to determine C_w, the derivative of the incumbents' objective function $\bar{\mathcal{U}}_\mathrm{p}$ with respect to p is given by [A.2.14] and is thus strictly decreasing for $p > \sqrt{p^\dagger}\sqrt{\pi(s)}$. Since $p^* > \sqrt{p^\dagger}\sqrt{\pi(s)}$, the proposed deviation is both feasible and payoff-improving for the incumbents. Hence, there is no equilibrium with this configuration of binding constraints.

The post-investment stage no-rebellion constraint involving only workers cannot bind in conjunction with a pre-investment stage constraint involving some incumbents

Suppose the pre-investment stage no-rebellion constraints were ignored. This leads to the choice of $p^* = \sqrt{p^\dagger}\sqrt{\pi(s)}$ as seen earlier. Given that there is no violation of [A.2.3a], if the resulting average payoff $\bar{\mathcal{U}}_\mathrm{p}^*$ were equally distributed among incumbents then the pre-investment stage no-rebellion constraint would be satisfied for all possible configurations of the rebel army (since $\mathcal{U}_\mathrm{p}^* = \mathcal{U}_\mathrm{p}'$). This equal distribution of payoffs has no negative consequences either directly for the incumbent objective function or for the conjectured binding constraints, so $p^* = \sqrt{p^\dagger}\sqrt{\pi(s)}$ would also be the equilibrium taking into account other pre-investment stage no-rebellion constraints. However, this has been seen to lead to a violation of a post-investment stage no-rebellion constraint involving incumbents (the proof that this violation occurs requires no assumptions on the distribution of incumbent payoffs, so choosing a non-equal distribution would not help in avoiding the violation). Therefore this case does not correspond to the equilibrium combination of binding no-rebellion constraints.

The post-investment no-rebellion constraint involving only workers binds in conjunction with a post-investment constraint involving incumbents. In this case, there must be complete payoff equalization among the group in power. All other no-rebellion constraints will hold.

Since some no-rebellion constraints must bind, the only remaining cases are ones where [A.2.3b] binds along with a post-investment stage no-rebellion constraint involving incumbents.

The binding version of [A.2.3b] is:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{w}} = \delta \frac{p}{p^{\dagger}}.$$

Since there is no shortage of workers to fill a rebel army, a post-investment stage no-rebellion constraint with a positive measure of incumbents can bind if and only if

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \underline{\mathcal{U}}_{\mathrm{p}} = \delta \left(\frac{p - p^{\dagger}}{p^{\dagger}} \right),$$

where $\underline{\mathcal{U}}_{\mathrm{p}}$ is the minimum among the payoffs received by any positive measure of incumbents. Note that the right-hand side must be strictly positive because $p > p^{\dagger}$. The argument is that if the left-hand side were greater than the right-hand side, the no-rebellion constraint would be violated since the constraint for workers is exactly binding with each exerting effort $\delta p/p^{\dagger}$, so if a positive measure of incumbents were to contribute power-adjusted fighting strength greater than $\delta p/p^{\dagger}$ per person then a successful rebellion would occur. On the other hand, if the left-hand side were strictly greater than the right-hand side then all positive measures of incumbents who are willing to rebel would contribute power-adjusted fighting strength strictly less than $\delta p/p^{\dagger}$. Since the no-rebellion constraint holds with equality for workers who each exert effort $\delta p/p^{\dagger}$, this means the no-rebellion constraint involving any positive measure of incumbents would be slack, contrary to our assumption.

Now suppose there is payoff inequality among a positive measure of incumbents. Given the definition of $\underline{\mathcal{U}}_{\mathrm{p}}$, it follows that $\underline{\mathcal{U}}_{\mathrm{p}} < \bar{\mathcal{U}}_{\mathrm{p}}$ because $\bar{\mathcal{U}}_{\mathrm{p}}$ is the average payoff. Equalization of payoffs would imply all incumbents receive $\bar{\mathcal{U}}_{\mathrm{p}}$, which given the inequality above, is such that:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \bar{\mathcal{U}}_{\mathrm{p}} < \delta \left(\frac{p - p^{\dagger}}{p^{\dagger}} \right).$$

But this implies every incumbent would bring power-adjusted fighting strength strictly less than $\delta p/p^{\dagger}$ when included in a rebel army. Since the constraint for workers is binding with fighting effort $\delta p/p^{\dagger}$, this means the post-investment stage no-rebellion constraint is now slack if any positive measure of incumbents is included. The equalization of payoffs has no direct effect on the incumbents' objective function and is seen to slacken a constraint that is binding, hence payoff inequality cannot occur in equilibrium.

Now suppose the post-investment no-rebellion constraint is binding for some incumbents in the case where there is payoff equalization among all positive measures of incumbents. Since the post-investment constraint for workers is binding, and as $p > p^{\dagger}$ implies there is no shortage of incumbents to fill a rebel army (and all are willing to rebel since $\mathcal{U}_{\mathrm{p}} < \mathcal{U}_{\mathrm{p}}^{\dagger}(K)$), because of the linearity of total effective fighting strength in the composition of the rebel army, this case is equivalent to the no-rebellion constraint binding for a rebel army including only incumbents:

$$\mathcal{U}_{\mathrm{p}}^{\dagger}(K) - \mathcal{U}_{\mathrm{p}} = \delta \left(\frac{p - p^{\dagger}}{p^{\dagger}} \right). \qquad \text{[A.2.25]}$$

Given that the post-investment no-rebellion constraint for workers is binding, the inequality [A.2.15] is equivalent to $p \geq \pi(s)$. Therefore, it must be the case that $p = \pi(s)$.

Now consider an equilibrium p^* and s^* where these values maximize $\mathcal{U}_{\mathrm{p}}$ subject to the power-sharing constraint (with consumption for workers determined by their binding no-rebellion constraint). Note that both binding no-rebellion constraints imply $\mathcal{U}_{\mathrm{p}}^* = \mathcal{U}_{\mathrm{w}}^* + \delta$. Using [3.6], it follows that $\mathcal{U}_{\mathrm{n}}^* = \mathcal{U}_{\mathrm{p}}^* - \delta + \alpha \mathcal{S}_{\mathrm{i}}(\tilde{\theta})$, and therefore:

$$\mathcal{U}_{\mathrm{p}}^* - \mathcal{U}_{\mathrm{n}}^* = \delta - \alpha \mathcal{S}_{\mathrm{i}}(\tilde{\theta}).$$

Since $\alpha \mathcal{S}_{\mathrm{i}}(\tilde{\theta}) \geq 0$, it follows that $\mathcal{U}_{\mathrm{p}}^* - \mathcal{U}_{\mathrm{n}}^* \leq \delta$, so [A.2.3a] is satisfied. As all incumbents receive the same payoff $\mathcal{U}_{\mathrm{p}}^{*} = \mathcal{U}_{\mathrm{p}}'$, all compositions of rebel armies at the pre-investment stage must satisfy their no-rebellion constraints. Therefore this case represents the combination of binding no-rebellion constraints in equilibrium.

The incumbents' payoff given the binding constraints

Given that both [A.2.8] and [A.2.25] hold, it follows that

$$\mathcal{U}_{\mathrm{w}} = \mathcal{U}_{\mathrm{p}} - \delta.$$

Substituting this expression for $\mathcal{U}_{\mathrm{w}}$ into the incumbents' objective function $\bar{\mathcal{U}}_{\mathrm{p}}$ from equation [A.2.2], and with payoff equalization $\bar{\mathcal{U}}_{\mathrm{p}} = \mathcal{U}_{\mathrm{p}}$:

$$p\mathcal{U}_{\mathrm{p}} = (1-p)(\mathrm{q} - (\mathcal{U}_{\mathrm{p}} - \delta)) + \mu(\kappa - \tilde{\theta})s,$$

and hence by rearranging the above to find an expression for $\mathcal{U}_{\mathrm{p}}$:

$$\mathcal{U}_{\mathrm{p}} = (\mathrm{q} + \delta)(1-p) + \mu(\kappa - \tilde{\theta})s.$$

Now substituting for $p = \pi(s)$ using the formula for $\pi(s)$ from [A.2.13]:

$$\mathcal{U}_{\mathrm{p}} = (\mathrm{q} + \delta)(1 - p^{\dagger}) + \mu(\kappa - \tilde{\theta})s - \frac{\mu\tilde{\theta}s}{\delta}.$$

Simplifying this expression yields

$$\mathcal{U}_{\mathrm{p}} = (\mathrm{q} + \delta)(1 - p^{\dagger}) + \mu\left(\kappa - \frac{\tilde{\theta}}{p^{\dagger}}\right)s,$$

and by substituting the formula for $p^{\dagger}$ from [3.5], the expression for $\mathcal{U}_{\mathrm{p}}$ in [3.7d] is obtained. This completes the proof.

A.3 Proof of Proposition 3

First, note that the relationship between s and $\tilde{\theta}$ in [3.7a] implies

$$\tilde{\theta} = \psi + (\kappa - \psi)s. \qquad [A.3.1]$$

Proposition 2 shows that the two constraints in [3.7b] are binding. By subtracting these equations from one another, it follows that:

$$\mathcal{U}_{\mathrm{w}} = \mathcal{U}_{\mathrm{p}} - \delta. \qquad [A.3.2]$$

The level of investment in equilibrium

By substituting the formula for $\tilde{\theta}$ from [A.3.1] into the expression for the incumbents' payoff in [3.7d]:

$$\mathcal{U}_{\mathrm{p}} = \frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta} + \mu\left(\kappa - \left(\frac{\mathrm{q}+2\delta}{\delta}\right)\psi - \left(\frac{\mathrm{q}+2\delta}{\delta}\right)(\kappa-\psi)s\right)s.$$

The derivative with respect to s is

$$\frac{\partial\mathcal{U}_{\mathrm{p}}}{\partial s} = \mu\left(\kappa - \left(\frac{\mathrm{q}+2\delta}{\delta}\right)\psi - 2\left(\frac{\mathrm{q}+2\delta}{\delta}\right)(\kappa-\psi)s\right).$$

Setting the derivative to zero and solving for s yields

$$s = \frac{\delta\kappa - (q+2\delta)\psi}{2(q+2\delta)(\kappa-\psi)} = \frac{1}{2}\frac{\delta\kappa - (q+2\delta)\psi}{(q+2\delta)\kappa - (q+2\delta)\psi}. \qquad \text{[A.3.3]}$$

Since $q+2\delta > \delta$, this expression can never be more than 1, but could be negative. Given that s is restricted to $s \in [0,1]$, the value of s that maximizes $\mathcal{U}_\mathrm{p}$ is

$$s^* = \max\left\{0, \frac{\delta\kappa - (q+2\delta)\psi}{2(q+2\delta)(\kappa-\psi)}\right\}. \qquad \text{[A.3.4]}$$

This is the solution given in [3.11]. It can be seen that s^* is positive whenever $\delta\kappa > (q+2\delta)\psi$, which is equivalent to $\kappa/\psi - 1 > 1 + q/\delta$.

To confirm that this is indeed an equilibrium, it is necessary to check whether several auxiliary constraints are satisfied. First, there is the condition from [2.5] that an incumbent would find it rational to defend the existing rules if there is a rebellion where the individual is not included in the rebel army. This requires $\mathcal{U}_\mathrm{p}^* > \mathcal{U}_\mathrm{n}' = \mathcal{U}_\mathrm{n}^*$ at the pre-investment stage, and $\mathcal{U}_\mathrm{p}^* > \mathcal{U}_\mathrm{w}^\dagger(K)$ at the post-investment stage.

Note that [3.6] implies the following expression for $\mathcal{U}_\mathrm{n}^*$:

$$\mathcal{U}_\mathrm{n}^* = \mathcal{U}_\mathrm{w}^* + \alpha\mathcal{S}_\mathrm{i}(\tilde{\theta}^*).$$

Given that $\mathcal{U}_\mathrm{p}^* = \mathcal{U}_\mathrm{w}^* + \delta$ according to [A.3.2], the condition $\mathcal{U}_\mathrm{p}^* > \mathcal{U}_\mathrm{n}^*$ is equivalent to

$$\alpha\mathcal{S}_\mathrm{i}(\tilde{\theta}^*) < \delta. \qquad \text{[A.3.5]}$$

Using [3.7a] and the distribution of the effort cost θ from [3.1], it is known that $0 < \psi < \kappa$ and $\psi \le \tilde{\theta}^* \le \kappa$. Hence it must be the case that $\tilde{\theta}^* - \theta < \kappa$ for all $\theta \in [\psi, \kappa]$, and the definition of $\mathcal{S}_\mathrm{i}(\tilde{\theta})$ in [A.2.1] can then be used to deduce $\mathcal{S}_\mathrm{i}(\tilde{\theta}^*) < \kappa$. Since α is a probability and $\kappa < \delta$ according to the parameter restrictions in [3.3], the condition [A.3.5] is verified.

Using the expression for $\mathcal{U}_\mathrm{p}$ from [3.7d], the expression for $\mathcal{U}_\mathrm{w}^\dagger(K)$ from [3.5], and the formula for the capital stock K, the condition $\mathcal{U}_\mathrm{p}^* > \mathcal{U}_\mathrm{w}^\dagger(K)$ is equivalent to:

$$\frac{(q+\delta)^2}{q+2\delta} + \mu\left(\kappa - \left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^*\right)s^* > \frac{(q+\delta)^2}{q+2\delta} - \delta + \mu\kappa s^*.$$

After cancelling terms and rearranging, this requirement reduces to:

$$\mu\left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^* s^* < \delta. \qquad \text{[A.3.6]}$$

Consider an equilibrium where $s^* > 0$, in which case equation [A.3.4] implies

$$(\kappa - \psi)s^* = \frac{\delta\kappa - (q+2\delta)\psi}{2(q+2\delta)},$$

and where it must be the case that $\delta\kappa > (q+2\delta)\psi$. Substituting this into the formula for $\tilde{\theta}$ in [A.3.1]:

$$\tilde{\theta}^* = \psi + \frac{\delta\kappa - (q+2\delta)\psi}{2(q+2\delta)} = \frac{1}{2}\left(\left(\frac{\delta}{q+2\delta}\right)\kappa + \psi\right).$$

Given that $\psi < \kappa$, it must be the case that $\psi < (\delta/(q+2\delta))\kappa$, and so the expression above for $\tilde{\theta}^*$ implies:

$$\tilde{\theta}^* < \left(\frac{\delta}{q+2\delta}\right)\kappa.$$

Since $\mu \le 1$ and $s^* \le 1$, and using the parameter restriction $\kappa < \delta$ from [3.3], the inequality above implies that [A.3.6] must hold, demonstrating that $\mathcal{U}_{\mathrm{p}}^{*} > \mathcal{U}_{\mathrm{w}}^{\dagger}(K)$.

As the utility function is linear, the link between worker and incumbent payoffs in [A.3.2] implies $C_{\mathrm{p}}^{*} = C_{\mathrm{w}}^{*} + \delta$. Therefore, all non-negativity constraints on consumption will hold if $C_{\mathrm{w}}^{*} \geq 0$, which is equivalent to $U_{\mathrm{w}}^{*} \geq 0$. Observe first that $s = 0$ is always a feasible choice for the incumbents in maximizing $\mathcal{U}_{\mathrm{p}}$, so if $s^{*} > 0$, it follows from the expression in [3.7d] that:

$$\mu\left(\kappa - \left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^{*}\right)s^{*} \geq 0. \qquad [A.3.7]$$

Substituting the expression for $\mathcal{U}_{\mathrm{p}}$ from [3.7d] into [A.3.2] yields:

$$\mathcal{U}_{\mathrm{w}}^{*} = \left(\frac{(q+\delta)^2}{q+2\delta} - \delta\right) + \mu\left(\kappa - \left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^{*}\right)s^{*}, \qquad [A.3.8]$$

and since [A.3.7] shows the second term is non-negative, a sufficient condition for $\mathcal{U}_{\mathrm{w}}^{*} \geq 0$ is

$$\frac{(q+\delta)^2}{q+2\delta} \geq \delta. \qquad [A.3.9]$$

The parameter restriction $\delta/q \leq \varphi$ from [3.3] implies that this inequality holds, so it is confirmed that $C_{\mathrm{w}}^{*} \geq 0$. Therefore, the solution in [A.3.4] is shown to be the unique equilibrium.

The constrained efficient level of investment

Using the relationship between p and s from [3.7c], it follows that:

$$\delta(1-p) = \delta\left(1 - p^{\dagger} - \frac{\mu\tilde{\theta}s}{\delta}\right) = \delta(1-p^{\dagger}) - \mu\tilde{\theta}s = \frac{\delta(q+\delta)}{q+2\delta} - \mu\tilde{\theta}s,$$

where the formula for $p^{\dagger}$ from [3.5] is also substituted into the above expression. The definition of the average payoff $\bar{\mathcal{U}}$ from [3.9] is equivalent to [3.10], and the expression for $(1-\delta)p$ above can be used to obtain:

$$\bar{\mathcal{U}} = \frac{(q+\delta)^2}{q+2\delta} - \frac{\delta(q+\delta)}{q+2\delta} + \mu\left(\kappa - \left(\frac{q+\delta}{\delta}\right)\tilde{\theta}\right)s + \mu\tilde{\theta}s + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}).$$

After simplification, this expression for $\bar{\mathcal{U}}$ reduces to:

$$\bar{\mathcal{U}} = \frac{q(q+\delta)}{q+2\delta} + \mu\left(\kappa - \left(\frac{q+\delta}{\delta}\right)\tilde{\theta}\right)s + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}).$$

Equation [A.3.1] gives a relationship between $\tilde{\theta}$ and s, which can also be substituted into the above:

$$\bar{\mathcal{U}} = \frac{q(q+\delta)}{q+2\delta} + \mu\left(\kappa - \left(\frac{q+\delta}{\delta}\right) - \left(\frac{q+\delta}{\delta}\right)(\kappa-\psi)s\right)s + \mu\mathcal{S}_{\mathrm{i}}(\tilde{\theta}). \qquad [A.3.10]$$

Using [3.1] and [A.2.1], an explicit expression for the expected surplus $\mathcal{S}_{\mathrm{i}}(\tilde{\theta})$ from receiving an investment opportunity is given by:

$$\mathcal{S}_{\mathrm{i}}(\tilde{\theta}) = \int_{\theta=\psi}^{\tilde{\theta}} \frac{\tilde{\theta}-\theta}{\kappa-\psi}\mathrm{d}\theta = \frac{1}{2}\frac{(\tilde{\theta}-\psi)^2}{\kappa-\psi} = \frac{1}{2}(\kappa-\psi)s^2,$$

where equation [3.7a] has been used to write this solely in terms of s. This is then substituted into [A.3.10] to obtain an expression for $\bar{\mathcal{U}}$ in terms of s:

$$\bar{\mathcal{U}} = \frac{q(q+\delta)}{q+2\delta} + \mu\left(\kappa - \left(\frac{q+\delta}{\delta}\right) - \left(\frac{2q+\delta}{2\delta}\right)(\kappa-\psi)s\right)s. \qquad [A.3.11]$$

Using [A.3.11], the derivative of $\bar{\mathcal{U}}$ with respect to s is:

$$\frac{\partial \bar{\mathcal{U}}}{\partial s} = \mu\left(\kappa - \left(\frac{q+\delta}{\delta}\right) - \left(\frac{2q+\delta}{\delta}\right)(\kappa - \psi)s\right).$$

Setting the derivative to zero and solving for s yields:

$$s = \frac{\delta\kappa - (q+\delta)\psi}{(2q+\delta)(\kappa-\psi)} = \frac{\delta\kappa - (q+\delta)\psi}{(q+\delta)\kappa - (q+\delta)\psi + q(\kappa-\psi)}. \quad \text{[A.3.12]}$$

Since $q+\delta > \delta$ and $\kappa > \psi$, this expression can never be greater than 1, but it could be negative. Therefore, if no auxiliary constraints are violated, the constrained efficient level of s is

$$s^{\diamond} = \max\left\{0, \frac{\delta\kappa - (q+\delta)\psi}{(2q+\delta)(\kappa-\psi)}\right\}. \quad \text{[A.3.13]}$$

This is the expression for $s^{\diamond}$ from [3.11]. It is positive whenever $\delta\kappa > (q+\delta)\psi$, which is equivalent to $\kappa/\psi - 1 > q/\delta$.

The auxiliary constraints to verify are $\mathcal{U}_{p}^{\diamond} > \mathcal{U}_{n}^{\diamond}$, $\mathcal{U}_{p}^{\diamond} > \mathcal{U}_{w}^{\dagger}(K)$, and $C_{w}^{\diamond} \geq 0$. Given that there is the same configuration of binding no-rebellion constraints, the analysis leading to [A.3.5] also shows that $\mathcal{U}_{p}^{\diamond} > \mathcal{U}_{n}^{\diamond}$ is equivalent to $\alpha\mathcal{S}_{i}(\tilde{\theta}^{\diamond}) < \delta$. Under the parameter restrictions from [3.3], this condition is necessarily satisfied.

Next, consider the constraint $\mathcal{U}_{p}^{\diamond} > \mathcal{U}_{w}^{\dagger}(K)$. Again, given that the configuration of binding no-rebellion constraints is the same, the analysis leading to [A.3.6] also applies in this case, so the requirement is equivalent to

$$\mu\left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^{\diamond}s^{\diamond} < \delta. \quad \text{[A.3.14]}$$

Consider a case where $s^{\diamond} > 0$. Using [A.3.13], it must then be the case that

$$(\kappa-\psi)s^{\diamond} = \frac{\delta\kappa - (q+\delta)\psi}{(2q+\delta)},$$

noting $\delta\kappa > (q+\delta)\psi$ is necessary. Substituting this into [A.3.1] yields:

$$\tilde{\theta}^{\diamond} = \psi + \frac{\delta\kappa - (q+\delta)\psi}{(2q+\delta)} = \frac{q\psi + \delta\kappa}{2q+\delta}. \quad \text{[A.3.15]}$$

Since $\psi < (\delta/(q+\delta))\kappa$ in this case, it follows that

$$\tilde{\theta}^{\diamond} < \frac{q\left(\frac{\delta}{q+\delta}\right)\kappa + \delta\kappa}{2q+\delta} = \frac{\delta(2q+\delta)\kappa}{(q+\delta)(2q+\delta)} = \frac{\delta}{q+\delta}\kappa,$$

which implies:

$$\left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^{\diamond} < \frac{q+2\delta}{q+\delta}\kappa.$$

Using the parameter restrictions in [3.3] and the inequality above:

$$\mu\left(\frac{q+2\delta}{\delta}\right)\tilde{\theta}^{\diamond} < \left(\frac{q}{2(q+2\delta)}\right)\left(\frac{q+2\delta}{q+\delta}\right)\kappa = \frac{1}{2}\frac{q}{q+\delta}\kappa < \kappa < \delta.$$

This demonstrates that [A.3.14] holds, which confirms that $\mathcal{U}_{p}^{\diamond} > \mathcal{U}_{w}^{\dagger}(K)$.

Finally, consider the non-negativity constraint $C_{w}^{\diamond} \geq 0$ for workers, which is equivalent to $\mathcal{U}_{w}^{\diamond} \geq 0$ given the utility function. Substituting the formula for $\mathcal{U}_{\mathrm{p}}$ from [3.7d] into [A.3.2], and using [A.3.15] to obtain an expression for $\tilde{\theta}^{\diamond}$:

$$\mathcal{U}_{\mathrm{w}}^{\diamond} = \left(\frac{(\mathrm{q}+\delta)^2}{\mathrm{q}+2\delta} - \delta \right) + \frac{\mu}{\delta(\mathrm{q}+2\delta)} \left(\delta(\mathrm{q}-\delta)\kappa - \mathrm{q}(\mathrm{q}+2\delta)\psi \right) s^{\diamond}. \qquad \text{[A.3.16]}$$

The sign of this expression is ambiguous for general parameters satisfying the restrictions in [3.3], so the non-negativity constraint for workers could be binding. When this expression is non-negative, the constrained efficient level of s is indeed given by the formula in [A.3.13] since all other auxiliary constraints are satisfied. More generally, since the non-negativity constraint is satisfied at $s = 0$, the possibility that it might be binding in equilibrium implies that $\kappa/\psi - 1 > \mathrm{q}/\delta$ is only a necessary condition for $s^{\diamond} > 0$.

Consider an equilibrium with $s^{\diamond} > 0$. In the case where the non-negativity constraint is not binding, the value of $s^{\diamond}$ is given by [A.3.13]. Comparison with the expression for s^* in [A.3.4] shows that $s^* < s^{\diamond}$. Now suppose the non-negativity constraint is binding. Since the non-negativity constraint is known to be satisfied at $s = 0$ and $s = s^*$, it follows that constrained efficient value of s must be strictly larger than s^*. Therefore, it is shown that $s^* < s^{\diamond}$ whenever $s^{\diamond} > 0$. This completes the proof.