Autocrats, Winning Coalitions, and Mass Revolts

Can the Public Tame a Dictator?

Thomas Apolte
University of Münster
Center for Interdisciplinary Economics
Scharnhorststrasse 100
48151 Münster
Germany
Apolte@wwu.de

October 4, 2016

Abstract

We link a simple threshold model of expressive behavior to an agency theory to explain how mass revolts may impact on a winning coalition's incentives to remain loyal to an incumbent dictator. Having observed public policy and updated their belief on the type of the government, the winning coalition's members may exploit the incident of a public revolt for escaping a loyalty trap that otherwise prevented them from switching to disloyalty. Our model implies an interacting effect of public revolts on the filters that obscure the relationship between defecting autocrats and the withdrawal of loyalty by the winning coalition in an autocracy. This interacting effect explains a rich set of empirical observations.

JEL classification: D02, H11, D74

Keywords: Accountability; Autocracy; Coups; Revolutions.

1 Introduction

An autocrat's power could effectively be contained to the extent that there were a credible threat of a potential overthrow that became effective whenever the autocrat pursued a non-welcomed public policy. Such a mechanism to be broadly beneficial requires some causal relationship between such a threat on the one hand and the wealth of a broad public on the other (Besley and Kudamatsu, 2008). The most direct causality is that from low public wealth to a "public rising", but this causality is plagued by the collective-action problem of revolutions (Tullock, 1971). This notwithstanding, rebellious activities by a broad public have at least been associated with a number of major challenges of political regimes, so that the real world does not always seem to fit into the picture of collective-action theory (see Kurrild-Klitgaard, 2004; Lichbach, 1998). This applies not only to the revolts against the communist regimes at the end of the 1980s but also to those of the so called Arab rebellion, and we have seen further public revolts in Thailand, Iran and the Ukraine, to name but a few. Understanding the potential of public control in autocracies presupposes this gap between theory and reality to be closed, and that is what this paper aims at contributing to.

Different branches of literature have developed ways to deal with the problem of collective action in public revolts, but as of yet, none of them is fully convincing. For example, the general equilibrium approach to insurgencies by Grossman (1991; 1999) can do without bypassing the collective-action problem. However, Grossman's approach is not exactly applicable to the—at leat seemingly—spontaneous outbreak of public revolts but rather to the formation of groups like *Hamas*, *Hizbollah*, or *ISIS*, that is with company-like organizations operating with a long-term perspective. Hence, they do not explain revolutionary events like those of 1989 in Middle- and Eastern Europe. By contrast, although the deprivation literature (Gurr, 1970) attracted considerable attention (Acemoglu and Robinson, 2001; Bloch, 1986; Boix, 2003), it ignores the problem or simply assumes the potentially revolting groups to somehow find ways for solving their collective-action problem

(Acemoglu and Robinson, 2006, pp. 126-128).

The more recent literature on selectorates defines a winning coalition by its capability to decide as to whether a government will stay in office or not. It does, however, not delve into the particular mechanisms that may back or oust a government (see Bueno de Mesquita, Smith, et al., 2005; Bueno de Mesquita and Smith, 2010; Besley, 2007; Besley and Kudamatsu, 2008). Should a broader public be capable of contributing to an overthrow of a government, then it is by definition part of a winning coalition. However, that naturally leaves the question unresolved as to whether there are further members of the winning coalition and, if so, how the different subgroups might interact in ousting a government. With the exception of Gilli and Li (2014; 2015), selectorate theory has so far not aimed at considering these questions.

Similar to the selectorate approaches, Myerson (2008) analyzes the credibility problems between a dictator and a winning coalition that either backs or, for that matter, deposes a dictator. Svolik (2009) models a Basian power game between the dictator and the winning coalition in which the latter might be able to deter a dictator from diverting from an implicit power-sharing contract with the ruling coalition. None of these approaches relate a violent overthrow of a dictator to the general public, however. They can hence explain coups but not public revolts and their impact on the power of a dictator.

By contrast, threshold models of collective behavior (see Granovetter, 1978; Schelling, 1978; Kuran, 1989; Marwell and Oliver, 1993; Lohmann, 1993; 1994, Yin, 1998) are indeed able to consistently explain public revolts. However, these approaches remain silent on the causality between mass protests on the one hand and the enforced resignation of an incumbent dictator on the other. Somewhat roughly speaking one may ask: Why should an incumbent step down "only" because there are public protests or even violent rebellions as long as he is backed by the winning coalition on which his power rests after all? Furthermore, one would ask why a winning coalition should quit backing the incumbent "only" because there are subgroups of the population protesting against the dictator? The winning coalition in a dictatorship typically comprises, *inter alia*, high officials of the police, the army, and the security forces, that is those that have the guns. And indeed, these officials have at times loyally executed orders by the political leaders to shoot at the protesters, like in Beijing 1989, but sometimes they either refused to do so or they even openly withdrew their loyalty from the political leadership. Only months after Beijing's security forces had committed the Tianmen massacre, the GDR's security forces refused to violently suppress the ongoing protests in Leipzig and East-Berlin and thereby set the stage for the spectacular collapse of a regime.

Two recent papers aim at relating public revolts to the threat of a coup d'etat. Svolik (2013) analyze the probability of a military coup against a certain threat of social unrest in light of the resources the government presumably needs the military to endow with for suppressing the public threat. Intermediate levels of the public threat come with highest coup probabilities since they combine a relatively strong military with only moderate incentives of the government to meet the military's political demands. The approach can explain correlations between public unrest and coups, but it assumes the intensity of the public threat as exogenous and hence as not related to the government's policy. This leaves it open as to whether the overthrow of an autocrat can be traced back to the deprivation of the public via public unrest.

Casper and Tyson (2014) define a military elite as well as a group of citizens both members of which want the government to be either ousted or, in the case of the citizens, urged to political reforms. Each military member's best response to a sufficient number of further military members having decided to rebel is to rebel, too, which applies to the group of the citizens in a likewise manner. As reform resistance and coup resistance of the government are assumed to be correlated, the observation of citizen protests are an informative signal to the elite members of a weak government and can thus serve as a focal point for the military members to coordinate on coup activities. However, since the collective-action problem of revolts are assumed to be solved by special benefits (Casper and Tyson, 2014, p. 552), the approach resembles that by (Grossman, 1991) and hence cannot account for spontaneous outbreaks of rebellions.

In this paper, we provide for an alternative link between collective action of a broader public on the one hand and the loyalty of a winning coalition to an autocrat on the other. On the basis of the presented model we can, under due consideration of the collective-action problem, explain why mass revolts, if they occur, sometimes sweep a dictator out of office and sometimes not. It implies that a dictator's public policy may be checked to some extent by the threat of mass revolts, although this threat is likely to be weak in real-world terms.

The paper is organized as follows. In section 2, we define all groups and subgroups of our model society on which our analysis rests throughout the paper. In section 3, we lay out a simple threshold model of expressive behavior. In section 4 we link the threshold model to an agency model of the relation between the incumbent and the winning coalition. In section 5, we summarize central findings as well as empirical implications, and in section 6, we conclude.

2 Structure of the Model Society

We lean on selecorate approaches in the definition of the groups and subgroups considered in our model society (see Bueno de Mesquita, Smith, et al., 2005; Bueno de Mesquita and Smith, 2009; Bueno de Mesquita and Smith, 2010; Besley and Kudamatsu, 2008). For the sake of our topic, however, we restrict the range of political regimes to autocracies of some sort.

First of all, we have the group of the entire population GP consisting of P domestic inhabitants. A subgroup $GS \subset GP$ comprising S members is referred to as the selectorate. This group is formally or informally endowed with the right to appoint the government. However, within the selectorate, only a subgroup is indeed decisive with respect to the recruitment of the government. As in Bueno de Mesquita, Smith, et al. (2005, pp. 51 - 55), we refer to this subgroup as the winning coalition $GW \subset GS$ which comprises W members. While basically consisting of a majority of the selectorate members, the winning coalition in an autocracy consists of an inner circle of the bureaucracy and, most importantly, of leading officials of the police, the military and further security forces.

Whenever the winning coalition withdraws its loyalty from the dictator the existing power structure will collapse. In some sort of an open contest, then, a new government will be inaugurated along with a newly established winning coalition (see Besley and Kudamatsu, 2008). The latter is assumed although it will always, and almost by definition, be the army, the police and further security forces as such that exercise coercive power and hence will always remain decisive. This does not imply, however, that the personal positions in the respective organizations remain unchallenged, and in particular not so at the top levels. It is indeed very likely that regime collapses come along with cascade effects of the withdrawal of loyalty to the government within the winning coalition. We do not analyze these internal processes, though, but restrict our considerations to the assumption that a regime collapse shatters an existing winning coalition's internal power structure resulting in the winning coalition to be newly established.

We define a further group $GL \in \{GP \backslash GW\}$ the members of which may or may not belong to the selectorate, but if they do they had previously been overruled in the broad sense of the word by the winning coalition. In any case, they have no influence whatsoever on any policy decision. With reference to a famous article by Vaclav Havel (1985), we refer to this group as the *powerless*. The incentives of each member of group GL of the powerless to take political action are compatible with some sort of expressive behavior at best, for no single member can intentionally exert any influence on the behavior or the position of an incumbent. If the individual members happen to coordinate on collectively rebelling against the government, however, the group as a whole may nevertheless impact on the behavior of the winning coalition and thus indirectly on the power position of the incumbent. Hence, while being individually powerless but doing things that may aggregate to a collectively significant hazard to the incumbent, one might arguably speak of some collective power, which is what Vaclav Havel (1985) obviously had in mind.

Finally, we define the government R_k with $k \in \{G, B\}$ which is of either type G (good) or type B (bad). The government is drawn from any subgroup of the population and we assume, for simplicity, that it consists of only a single person.

3 Sparking Public Revolts

Let $g \in [0, 1]$ be the ratio of public-goods expenditures in terms of an exogenous level of tax revenues. We assume all excess tax revenues $1 - g$ to be either spent as transfers paid to the members of the winning coalition or to be retained and used for concealed private consumption by the government officials. Independently of the government's choice g, each member of group GL has a private opinion regarding the minimum of a public expenditure ratio that this particular member accepts as appropriate, given his or her evaluation of agency costs and possibly also some tolerated degree of governmental slack. We follow the deprivation literature (Gurr, 1970; Bloch, 1986; Boix, 2003; Acemoglu and Robinson, 2006) and define the individual degree of relative deprivation of any member i of group GL as $\gamma_i := \frac{g_i}{g} - 1$, where $g_i \in [g, 1)$ is the lowest public-expenditure ratio that member i would just be willing to accept.

For convenience, we define a group GL and normalize its size to unity. Leaning on Kuran (1989), we assume a share $z \in [0, 1]$ of group GL to exhibit a disobedient habit toward the government. Disobedience can take a range of different forms: It may be limited to statements or comments among friends or, within a more general public, it may imply the attendance in peaceful demonstrations; but it may as well go as far as to the participation in violent rebellious activities. In any case, however, the character of these individual activities is purely expressive (Brennan and Lomasky, 1993; Hillman, 2010) in that they do, from the point of view of the individual, not aim at increasing the probability of an overturn of the government—although they may effectively contribute to precisely that. The share z of disobedient members of group GL cannot directly be observed by the members of GL, which is particularly relevant in a dictatorship, where there is no free flow of information. Each member of group GL is hence reliant on an observation z^e that only imperfectly indicates the true share z, and only with a certain time lag. As is standard in threshold models, however, we abstract from the imperfection of the observation z^e but consider the time lag.

In the tradition of "classical" threshold models (see Granovetter, 1978; Kuran, 1989; Marwell and Oliver, 1993) we construct a cascade effect driven by the presumed impact of disobedient behavior by some members of group GL on the level of expressive utility from disobedient behavior of further members. This approach is different from those by Lohmann (1993; 1994) and Casper and Tyson (2014) in that the latter are constructed around informational cascades rather than around expressive utility. As compared to the expressive-utility approach the informational-cascade models impress with their capability of accounting for more complex empirical observations, in particular those around the fall of the Berlin wall in 1989 (Lohmann, 1994). However, these empirical fits are based on case-study material rather than on statistically significant relationships. A more fundamental issues is this: By the same token of the more sophisticated empirical patterns the informational cascades are reliant on rather strong instrumental incentives for individual protest participation. Lohmann (1993) refers to equilibrium participation rates in elections (Ledyard, 1984) for tackling the associated public-goods problem, but this bypasses that problem rather than convincingly solving it. At least for our purposes, hence, an expressive-utility approach is more convincing and it is fully applicable despite its somewhat less sophisticated empirical implications.

We assume the degree of relative deprivation to be distributed according to the following cumulative distribution function:

$$H(\gamma) = \frac{2a - \gamma}{\gamma} \quad with : H \in [0, 1] \quad and : a \geq 1, \tag{1}$$

which says that each member of group GL has a degree of relative deprivation of $\gamma = a$ or more. Since H rises in a for each level of γ, the parameter a is an indicator of the overall degree of relative deprivation within group GL. Consider now a marginal member m of group GL in the sense that this member is just indifferent between obedience and disobedience. Then $H(\gamma_m)$ gives the share z of disobedient members of GL so that $H(\gamma_m) = z$, which implies:

$$z = \frac{2a - \gamma_m}{\gamma_m}. \tag{2}$$

By implication, rising shares z of disobedient members correspond to dropping levels of relative deprivation γ_m of a marginal member. Next, we define a function that gives the level B of expressive utility as it stems from a disobedient habit of a marginal member of group GL toward the government:

$$B = \gamma_m z^e. \tag{3}$$

The underlying assumption here is that disobedience yields expressive utility only upon the observation that further members of one's peer-group members express disobedience too (see Hillman, 2010). The cascade effect hence rests on expressive utility externalities rather than on cost externalities. The former is more convenient for our purposes and empirically at least as plausible. Leaning on costs instead of expressive utility would, however, not change any of our results. Combining 2 and 3 yields:

$$B = \frac{2az^e}{1 + z}. \tag{4}$$

By assumption, the government can impose a cost $c \in [0, a]$ on each disloyal member of group GL as far as the security forces assume the relevant governmental orders. Since a marginal member of GL is defined over $B = c$, we can substitute c for B in equation 4 and find, upon solving for z, the following "threshold" condition:

$$z = -1 + \frac{2a}{c} z^e. \tag{5}$$

Since we assumed z^e to be a somewhat lagged but otherwise perfect indication of the true value z, we can subtract z^e on both sides of the threshold condition 5 for reaching at the following dynamic version of the threshold function:

$$\dot{z} = -1 + \frac{2a - c}{c} z^e. \tag{6}$$

Following the literature on threshold models (Marwell and Oliver, 1993), we refer to the steady-state share $z = z^e$ of disobedient members of GL as the "critical share" z^e_{cr} . Substituting z^e_{cr} for z and z^e in the threshold condition 5 yields:

$$z^e_{cr} = \frac{c}{2a - c}. \tag{7}$$

Figure 1 illustrates the dynamics. Line $\dot{z}(z^e)$ represents equation 6. For any initial value of the expectation $z^e < z^e_{cr}$, the change $\dot{z}$ in disloyal members is negative which further reduces the expected level z^e until it reaches zero. We refer to $z^e = z = 0$ as a peace equilibrium, since there are no disobedient members of group GL of the powerless and hence no protests or other related activities. Should the expected level z^e ever exceed the threshold z^e_{cr}, then the change in disobedient members will be positive which further increases the expected level z^e until it reaches unity. We refer to $z^e = z = 1$ as a rebellion equilibrium, since all members of group GL are disobedient which implies activities like public protests or even violence.

We now introduce a shock that disturbs the expectation $z^e = 0$ in an es-

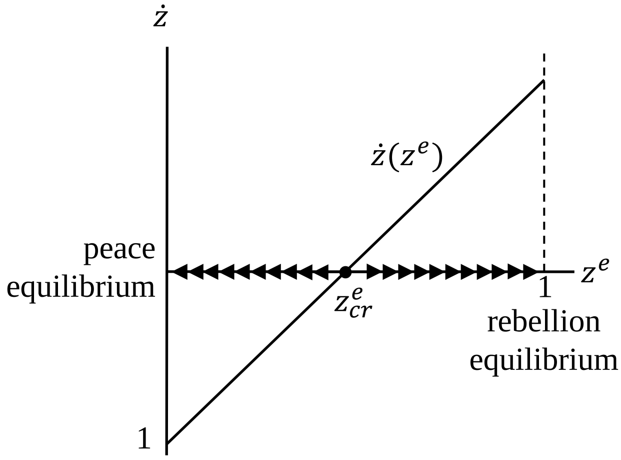

Figure 1: Dynamics of Insurrections

tablished peace equilibrium whenever new information on the tax allocation chosen by the government is conveyed.[1] The extent $u \in [0, 1]$ of such a shock follows a truncated probability density function $Tr(u)$ for any given set of new information on the tax allocation. Following the dissemination of the latter in an established peace equilibrium, the expected share of disobedient members of group GL is hence $z^e = u$. With $\rho := \int_{z^e_{cr}}^{1} Tr(u)\, \mathrm{d}u$, we define the probability that $u > z^e_{cr}$ and hence that the shock is sufficiently strong for turning a peace equilibrium into a rebellion equilibrium. For our analysis, we assume $Tr(u)$ to be common knowledge. However, this does not naturally fit reality, which will be discussed further below.

Since the critical value z^e_{cr} is a function of the parameter a as well as of the cost c, the probability ρ, too, is a function of these two parameters for any given distribution of stochastic shocks $Tr(u)$. In particular, since $z^{e'}_{cr}(a) < 0$ and $z^{e'}_{cr}(c) > 0$, a rise in the overall level of relative deprivation within group GL or a drop in the costs of disloyalty lower the critical value z^e_{cr} and thus raise the probability of a process toward a rebellion equilibrium following the dissemination of new information on the tax allocation, so that:

[1]We abstract from such shocks in a rebellion equilibrium.

$$\rho = \rho(a, c) \qquad with : \rho'(a) > 0;\ \rho'(c) < 0. \tag{8}$$

Note that a probability $\rho > 0$ of a rebellion equilibrium presupposes a threshold level $z^e_{cr} < 1$ which, according to equation 7, applies to all $c \in [0, a)$. By contrast, we have $z^e_{cr} = 1$ and hence $\rho = 0$ whenever $c = a$ so that, at least formally, the probability of a rebellion can be reduced to zero by raising the costs to a level a. At least in a dictatorship, the police, the military and the secret services are typically part of the winning coalition. Since it is the winning coalition that has the guns and thus control over the costs of disloyalty, it is, in principle, capable of raising the costs to a level necessary for rendering any dynamic into a rebellion equilibrium impossible.

This, however, raises the question as to why the winning coalition does not always and everywhere proceed in that way? Kuran (1989) argued that once the security forces have failed to raise the costs at an early stage, and once a rebellion equilibrium has settled, the security forces may simply shy away from raising the costs to such a tremendous level as is necessary for restoring a peace equilibrium. That would explain why the security forces abstained from shooting at the protesters in Warsaw, East-Berlin, Prague, and elsewhere in 1989. But it would then remain to be explained why the security forces in Beijing did indeed shoot.

Svolik's (2013) approach predicts a well endowed military or, more generally, well endowed security forces like that in China to oust the government rather than to remain loyal and shoot at the citizens. Casper and Tyson (2014), by contrast, assume not the endowment of the security forces but rather their capability of coordinating on a common coup strategy to be decisive. Still differently, one may also argue that the government in Beijing simply stood firm while those in Middle and Eastern Europe resigned in light of the protesting masses. However, the then head of the GDR government Erich Honecker showed no signs of resignation but rather released a formal command that stipulated shooting at the demonstrating masses (Lohmann, 1994, p. 69). The command induced preparations of the formal security forces in line with an armament of paramilitary forces which soon stood ready to strike. Surprisingly, though, they eventually declined assuming Honecker's command and thus opened the stage for the final act of the regime's demise.

So, why did the military elite shoot in Beijing but not so in Berlin? Our explanation is this: Whenever a rebellion equilibrium evolves, the members of the winning coalition evaluate their personal future prospects in the old or, alternatively, in a new regime against the background of their updated belief on the character of the incumbent government. Depending on this update, they either abstain from raising the protesters' costs and hence withdraw their loyalty; or they remain loyal to the incumbent and raise the costs of participation in a rebellion to a level sufficiently high as to suppress the rebellion. The parameters behind this decision will be different from case to case, and so will be the decision of the winning coalitions' members.

At the heart of the following considerations is the following: Combining a reticent habit toward raising the rebels' costs during an ongoing rebellion with withdrawing their loyalty can provide an opportunity for the winning coalitions' members to escape a loyalty trap in the sense of Bueno de Mesquita, Smith, et al. (2005) which binds them to even such an incumbent who is under the suspicion of cheating the winning coalition's members. In the following section, we lay out the details of our hypothesis.

4 Public Policy, Revolts, and Loyalty

4.1 Structure of the Model

As in the previous section, we assume the government to spend a share g of tax revenues for public goods. These goods are not only purely public in the *Samuelsonian* sense that no member of GP can be excluded from their consumption but also that they are not subject to any rivalry in consumption whatsoever. Apart from g, however, the government distributes a share vW of tax revenues as direct money transfers equally to each member of the winning coalition and the government.

Finally, the government may retain a share $e = 1 - g - vW$ of tax revenues and use it for government purposes alone. Funds e are not directly consumed by the government, nor are they direct transfers. Rather, they are used as inputs for the government sector in a way as to enhance the utility derived from holding a government position. This may as well imply losses due to governmental slack or the like. As e is no direct transfer, however, its disposability is low as compared to direct transfers, it is related to high positive externalities with respect to the incumbent's environment and, politically, allocating taxes into e becomes *ceteris paribus* the more delicate, the higher is e. The latter is particularly true when e is compared to direct transfers that are viewed as legitimate at least by the winning coalition. All in all, the utility derived from any unit of e is lower than what can be derived from direct transfers and, most notably, it is subject to substantially decreasing marginal utility. In any case, of course, the way these funds are used is not considered legitimate by members of both group GL of the powerless and the winning coalition GW. What is more, it is not even considered legitimate by a *good government* R_G. Hence, a good government will not be interested in retaining the share e in the first place since this would require the government to allocate the funds into non-legitimate channels. Whether or not a government is *good* or *bad* is not directly observable by either group GL or group GW, but the probability π of a government of being good is common knowledge.

We catch these aspects by describing indirect utility V^j as derived by groups $j \in \{GL, GW, R_k\}$ in the following functional form:

$$V^j = g^\alpha (1+v)^\phi e^\theta \; with \; \theta = \begin{cases} \beta \; for \; j = R_B \\ 0 \; for \; j = GL, GW, R_G \end{cases}$$

$$\phi = \begin{cases} 0 \ \ for\ group\ GL \\ 1 \ \ otherwise \end{cases} \qquad and\ \ 0 < \alpha, \beta < 1. \tag{9}$$

The budget constraint for public expenditures is:

$$1 = g + vW + e. \tag{10}$$

The respective group members find the optimal allocation of tax revenues (g_j^*, v_j^*, e_j^*) from their respective point of view by maximizing a group member's indirect utility, subject to the budget restriction 10. Table 1 summarizes the optimal allocations from the respective point of view of each group.[2] Note that $g_{R_k}^{*'}(W) > 0$ and $g_{GW}^{*'}(W) > 0$, indicating that expenditures for public goods become more attractive to both types of the government as well as for the winning coalition as the size of the winning coalition rises.

By the same token, we have $v_{GW}^{*'}(W) < 0$ and $v_{R_k}^{*'}(W) < 0$ since direct transfers to the members of the winning coalition become more expensive as the size of the winning coalition rises, which makes them less attractive relative to public goods. These results reproduce an implication from the selectorate model by Bueno de Mesquita, Smith, et al. (2005, pp. 77 - 106) as well as from Olson's encompassing-interest approach (Olson, 1993; McGuire and Olson, 1996). Both approaches imply that winning coalitions that grow in size will shift fiscal expenditure from redistribution in favor of privileged groups to the funding of public goods that are equally available to everybody.

Note further that this logic does not apply to e_{R_B} since the government does not need to share these funds with further members of the winning coalition.

[2] For the sake of brevity, we have assumed $g_{R_B}^*, v_{R_B}^*, e_{R_B}^* > 0$ as well as $g_{GW}^*, v_{GW}^* > 0$. Corner solutions $e_{R_B}^* = 0, v_{R_B}^* > 0; e_{R_B}^* > 0, v_{R_B}^* = 0; e_{R_B}^* = 0, v_{R_B}^* = 0$, and $g_{GW}^* = 0, v_{GW}^* = 0$ are possible (though not always plausible), but presenting all these cases would require lengthy considerations without adding further insights, nor would it change any of the results. A full set of Kuhn-Tucker conditions is of course available from the author.

Table 1: Optimal Tax Allocation

group	g	v	e
GL	$g^*_{GL} = 1$	$v^*_{GL} = 0$	$e^*_{GL} = 0$
GW or R_G	$g^*_{GW} = \frac{\alpha(1+W)}{1+\alpha}$	$v^*_{GW} = \frac{1-\alpha W}{(1+\alpha)W}$	$e^*_{GW} = 0$
R_B	$g^*_{R_B} = \frac{\alpha(1+W)}{1+\alpha+\beta}$	$v^*_{R_B} = \frac{1-(\alpha+\beta)W}{(1+\alpha+\beta)W}$	$e^*_{R_B} = \frac{\beta(1+W)}{1+\alpha+\beta}$

Hence, a rise in the size of the winning coalition makes these expenditures more attractive to a bad government, as can be seen by $e^{*'}_{R_B}(W) > 0$, which implies that a bad government tends to reallocate more funds away from bigger as compared to smaller winning coalitions. The rationale behind this is simply the rivalry in consumption of benefits to members of privileged groups. Finally, note that the winning coalition or a good government will always supply a higher level of public goods than a bad government since $g^*_{GW} > g^*_{R_B}$. Good governments will hence not only abstain from reserving tax revenues for own purposes, but they will also provide more public goods.

Table 2 presents the indirect utilities of groups GL and GW as well as of R_G and R_B as they are optimal from the point of view of the respective groups. As an example, if group R_B were decisive for the allocation of g_j, v_j, and e_j, then the allocation were $g^*_{R_B}, v^*_{R_B}$, and $e^*_{R_B}$, and the resulting indirect utility of group GL were $V^{GL}_{R_B} = g^{*^\alpha}_{R_B}$.

Remember that the winning coalition appoints the government and, further on, that it expects the government to allocate taxes in a way as to maximize the indirect utility of a winning coalition's member. Hence, a winning coalition's member wants the government to set g_j, v_j, and e_j such that its ensuing utility turns out to be V^{GW}_{GW}. However, a bad government may have a different plan. Recall that the winning coalition does not know in advance whether the government is good or bad. The powerless, in turn, would not be happy with a bad government's tax allocation either, since that implies less public goods compared to the level supplied by a good government. On

Table 2: Levels of utility of group... optimal for group ...

utility of... optimal for...	GL	GW, R_G	R_B
GL	$V^{GL}_{GL} = g^{*\alpha}_{GL}$	$V^{GW}_{GL} = g^{*\alpha}_{GL}$	$V^{R}_{GL} = g^{*\alpha}_{GL}$
GW, R_G	$V^{GL}_{GW} = g^{*\alpha}_{GW}$	$V^{GW}_{GW} =$ $g^{*\alpha}_{GW}(1 + v^{*}_{GW})$	$V^{R}_{GW} =$ $g^{*\alpha}_{GW}(1 + v^{*}_{GW})$
R_B	$V^{GL}_{R_B} = g^{*\alpha}_{R_B}$	$V^{GW}_{R_B} =$ $g^{*\alpha}_{R_B}(1 + v^{*}_{R_B})$	$V^{R}_{R_B} =$ $g^{*\alpha}_{R_B}(1 + v^{*}_{R_B})(1 + e^{*}_{R_B})^{\beta}$

top of that, a change in the attribution of individual members of the population to the respective subgroups GW, GL, and R_k, which may ensue from a broader change in government or even regime, would give each of the members of the powerless a chance for becoming member of a newly constituted winning coalition. In other words: The powerless would always win from a change in the power structure.

However, each member of the powerless is subject to the public-goods problem and does not face a sufficiently strong individual incentive for intentionally contributing to a change in the power structure. This notwithstanding, such a member might nevertheless be inclined to express his or her disapproval with a given tax allocation publicly and possibly even violently; but this inclination to translate into manifest collective action requires the powerless to get uncaged from their peace-equilibrium trap by an exogenous shock strong enough for shifting the initial expected value z^e of disloyal members of group GL above its critical level z^e_{cr}. As described in section 3, this does only happen with probability $\rho(a, c)$.[3]

[3]An interesting corollary of $\rho(a)' > 0$ in equation 8 is that processes toward a rebellion equilibrium are less likely in regimes with large winning coalitions and *vice versa*. The reason is again the encompassing-interest effect that induces governments backed by large winning coalitions to supply higher levels of public goods and thereby provide lower degrees of deprivation which, in turn, reduces the propensity of members of group GL to express

The winning coalition, in turn, would want to continue supporting the government as long as its members expect the latter to choose g^*_{GW}, v^*_{GW}, and e^*_{GW} in the future, and independently of the obedience or disobedience of members of group GL of the powerless. On the other hand, should the winning coalition expect the government to choose an allocation $g^*_{R_B}, v^*_{R_B}$, and $e^*_{R_B}$, then its members need to decide: If they accept this allocation, then their utility falls short of what it otherwise could have been. By contrast, should they decide to drop the government, then a new government will be appointed, but this new government will come along with a new winning coalition of which each individual will become a member only with probability W/S. The latter gives rise to a loyalty trap (Bueno de Mesquita, Smith, et al., 2005).

Should it happen, though, that new information on the government's choice of g sparks a process toward a rebellion equilibrium on the side of group GL of the powerless, then this may blaze a trail for the winning coalitions' members out of their loyalty trap. In particular, the winning coalition's members can jump on the protesters' bandwagon and support the ongoing rebellious activities by keeping the protesters' costs low. They may, for example, abstain from violence and from seriously prosecuting demonstrators, perhaps initially in a concealed way, and they may even proceed to actively support rebellious groups in various ways. Eventually, the winning coalition may openly turn against the incumbent dictator by formally withdrawing their loyalty. Doing so raises the probability of each member of the existing winning coalition of becoming a member of a newly established winning coalition above W/S and that is why a winning coalition may withdraw its loyalty to a government in light of ongoing public protests.

We explore the relation between the loyalty trap and mass revolts within a simple game that runs through two periods $t \in \{1, 2\}$. Players are nature (N), the period-one winning coalition (GW), and the period-one government (R_k). All members of both GW and R_k are drawn from the population GP. We assume all actors to be risk neutral. The game always starts in a peace

disobedient behavior. See a proof in appendix A.

equilibrium in period $t = 1$.

The government chooses a level $e_t \in \{0, e^*_{R_B}\}$ in each period t. In period $t = 1$ the information on e_t affects the expected level z^e of disobedient members within group GL of the powerless by an amount u. The level $e_t = 0$ is at least part of an optimal tax allocation not only from the perspective of the winning coalition and a good government but also from the perspective of the powerless, while a level $e_t = e^*_{R_B}$ is only optimal for a bad government. We capture this by assuming $u(e_1 = 0) = 0$ and $u(e_1 = e^*_{R_B}) \in (0, 1)$. As a consequence $\rho(e_t = 0) = 0$ and $\rho(e_t = e^*_{R_B}) \in (0, 1)$. In particular, the timing of the game is as follows:

1. Nature randomly selects a period-one winning coalition GW with probability W/S of each member of GP for being part of GW. Nature then randomly selects a period-one government R_k of type $k \in \{G, B\}$ from group GP, with probability π for $k = G$ and $1 - \pi$ for $k = B$.

2. The period-one government R_k chooses $e_1 \in \{0, e^*_{R_B}\}$ and period-one payoffs are realized.

3. If the government had chosen $e_1 = e^*_{R_B}$ in period 1, then nature decides at the beginning of period 2 with probability ρ that there will be a rebellion equilibrium and with probability $1 - \rho$ that there will be a peace equilibrium. If the government had chosen $e_1 = 0$, $\rho = 0$ and there will always be a peace equilibrium.

4. Group GW chooses among the options "support government" (SG) and "drop government" (DG).

5. The period-two winning coalition and the period-two government are determined depending on the winning coalition's choice between SG and DG in step 4:

   - If the winning coalition had chosen option SG in step 4, then the period-two winning coalition and the period-two government remain as they were in period one, independently of what the government had chosen in step 2 and independently of whether there is a peace equilibrium or a rebellion equilibrium.

- If the winning coalition had chosen option DG in step 4 in a peace equilibrium, then nature randomly selects a period-two winning coalition GW with probability W/S for each member of GP for becoming part of period-two GW. Nature then randomly selects a period-two government R_k of type $k \in \{G, B\}$ with probability π for $k = G$ and $1 - \pi$ for $k = B$ from group GP.
- If the winning coalition had chosen option DG in step 4 in a rebellion equilibrium, then the winning coalition remains as it was in period 1. Nature randomly selects a period-two government R_k of type $k \in \{G, B\}$ with probability π for $k = G$ and $1 - \pi$ for $k = B$ from group GP.

6. The period-two government R_k chooses $e_2 \in \{0, e^*_{R_B}\}$.
7. The period-two payoffs are realized and the game ends.

Two technical remarks are in order. Firstly, as we assume a large number of members of the total population GP we can safely neglect the expected value of additional future incomes of any member of GP for the case of this particular member to be appointed as period-two government. Secondly, since the winning coalition cannot extract any information on the type of the period-two government from the period-one government's decision on e_1 in the case of a change in government, it is left to calculate the expected value of e_2 by the prior probability π, so that $e_2 = (1 - \pi)e^*_{R_B}$ following a decision DG in step 4.

4.2 Equilibria

We solve for Perfect Bayes Equilibria based on the payoffs and definitions presented in table 3.[4] Depending on the parameters $k, \delta, \frac{W}{S}$, and ρ, we will either have a pooling equilibrium where any government chooses $e_1 = 0$, or

[4] Detailed proofs are presented in appendix B. See also a game tree in appendix C.

Table 3: Overview of Payoffs

$e_1 = e^*_{R_B}$	RDG	$SG/e_2 = 0$	$SG/e_2 = e^*_{R_B}$	PDG
	$U1$	$U2$	$U3$	$U4$
GW	$V^{GW}_{R_B} + V_{RDG}$	$V^{GW}_{R_B} + V^{GW}_{GW}$	$2V^{GW}_{R_B}$	$V^{GW}_{R_B} + V_{PDG}$
R_B	$V^{R}_{R_B} + \delta V_{LP}$	$V^{R}_{R_B} + \delta V^{R}_{GW}$	$(1+\delta)V^{R}_{R_B}$	$V^{R}_{R_B} + \delta V_{LP}$
R_G	$V^{GW}_{R_B} + \delta V_{LP}$	$V^{GW}_{R_B} + \delta V^{GW}_{GW}$	$(1+\delta)V^{GW}_{R_B}$	$V^{GW}_{R_B} + \delta V_{LP}$
$e_1 = 0$		$SG/e_2 = 0$	$SG/e_2 = e^*_{R_B}$	PDG
		$L1$	$L2$	$L3$
GW		$2V^{GW}_{GW}$	$V^{GW}_{GW} + V^{GW}_{R_B}$	$V^{GW}_{GW} + V_{PDG}$
R_B		$(1+\delta)V^{R}_{GW}$	$V^{R}_{GW} + \delta V^{R}_{R_B}$	$V^{R}_{GW} + \delta V_{LP}$
R_G		$(1+\delta)V^{GW}_{GW}$	$V^{GW}_{GW} + V^{GW}_{R_B}$	$V^{GW}_{GW} + \delta V_{LP}$

$V_{LP} := \pi V^{GL}_{GW} + (1-\pi)V^{GL}_{R_B}$ $\qquad$ $V_{RDG} := \pi V^{GW}_{GW} + (1-\pi)V^{GW}_{R_B}$

$$V_{PDG} := \tfrac{W}{S}\left[\pi V^{GW}_{GW} + (1-\pi)V^{GW}_{R_B}\right] + (1-\tfrac{W}{S})\left[\pi V^{GL}_{GW} + (1-\pi)V^{GL}_{R_B}\right]$$

RDG: Option *DG* within a rebellion equilibrium.
PDG: Option *DG* within a peace equilibrium.

a separating equilibrium with $e_1 = 0$ in the case of a good government and $e_1 = e^*_{R_B}$ in the case of a bad government.

If the government expects the only sufficiently safe way for preventing a mass rebellion followed by a coup to be $e_1 = 0$ and if it discounts the future not too much, then even a bad government will choose $e_1 = 0$. In that case, there will be a pooling equilibrium. We simplify notation by defining $A := (V^{GW}_{R_B} - V_{LP})/(\pi V^{GW}_{GW} + (1-\pi)V^{GW}_{R_B} - V_{LP})$ as well as $C := (V^{R}_{R_B} - V^{R}_{GW})/(V^{R}_{R_B} - V_{LP})$

and get:

Proposition 1. *For any* $C \leqslant \delta\rho|\frac{W}{S} \in [0, A)$ *or* $C \leqslant \delta|\frac{W}{S} \in [A, 1]$ *there is a pooling equilibrium in the following strategies:*

$$s^*_{GW} = \{SG\}, \tag{11}$$

$$s^*_{R_k} = \begin{cases} e_t = 0 \quad if \quad k = G; \\ e_1 = 0; e_2 = e^*_{R_B} \quad if \quad k = B \wedge C \leqslant \delta|\frac{W}{S} \in [A, 1]; \\ e_1 = 0; e_2 = e^*_{R_B} \quad if \quad k = B \wedge C \leqslant \delta\rho|\frac{W}{S} \in [0, A). \end{cases} \tag{12}$$

Proof: See appendix B.1 ■

For certain parameter values δ, ρ and $\frac{W}{S}$, however, the probability of the winning coalition to get freed from a loyalty trap by a public rebellion becomes sufficiently low as to make a choice $e^*_{R_B}$ worth the risk for a bad government. Alternatively, a bad government may discount the future so much that the payoff from $e^*_{R_B}$ exceeds any foregone period-two payoffs. In both cases, there will be a separating equilibrium:

Proposition 2. *For any* $C > \delta\rho|\frac{W}{S} \in [0, A)$ *or* $C > \delta|\frac{W}{S} \in [A, 1]$ *there is a separating equilibrium in the following strategies:*

$$s^*_{GW} = \begin{cases} SG \quad if \quad e_1 = 0; \\ DG \quad if \quad e_1 = e^*_{R_B} \wedge re; \\ DG \quad if \quad e_1 = e^*_{R_B} \wedge pe \wedge \frac{W}{S} \in [A, 1]; \\ SG \quad if \quad e_1 = e^*_{R_B} \wedge pe \wedge \frac{W}{S} \in [0, A), \end{cases} \tag{13}$$

$$s^*_{R_k} = \begin{cases} e_t = 0 \quad if \quad k = G; \\ e_t = e^*_{R_B} \quad if \quad k = B \wedge C > \delta|\frac{W}{S} \in [A, 1]; \\ e_t = e^*_{R_B} \quad if \quad k = B \wedge C > \delta\rho|\frac{W}{S} \in [0, A). \end{cases} \tag{14}$$

Proof: See appendix B.2 ■

Our two propositions imply a pooling equilibrium and three variants of a separating equilibrium, depending on the relevant parameters. All variants of the two possible equilibria are summarized in table 4.

Table 4: Perfect Bayes Equilibria

pooling equilibrium	**separating equilibrium**
$e_t = 0, SG, for\ k = G$ $e_1 = 0, e_2 = e^*_{R_B}, SG,\ for\ k = B$	$e_t = 0, SG, for\ k = G$ $e_1 = e^*_{R_B}, DG\ or\ SG, for\ k = B$
$C \leqslant \delta\rho \vert \frac{W}{S} \in [0, A)$ or $C \leqslant \delta \vert \frac{W}{S} \in [A, 1]$	loyalty trap $(1-\rho)$: $C > \delta\rho \vert \frac{W}{S} \in [0, A) \wedge pe$ accidental revolution (ρ): $C > \delta\rho \vert \frac{W}{S} \in [0, A) \wedge re$ hazarded coup: $C > \delta \vert \frac{W}{S} \in [A, 1] \wedge pe$

See figure 2 for the intuition behind the equilibria. The horizontal line on the left-hand side of the figure represents $C = \delta$, above which the conditions for a pooling equilibrium are satisfied for regimes characterized by $\frac{W}{S} \in [A, 1]$. By contrast, the hyperbolic line on the right-hand side represents $C = \delta\rho$, above which the conditions of a pooling equilibrium are satisfied for regimes characterized by $\frac{W}{S} \in [0, A)$.

Whenever there is a regime the size W/S of its winning coalition relative to the size of its selectorate is small, the probability of any person of the population of becoming part of a new winning coalition is low. If that relative size is as low as $\frac{W}{S} \in [0, A)$, the winning coalition is stuck in a loyalty trap unless a public rebellion happens to occur. This is the interesting case with respect to our topic, and it is depicted on the right-hand side of figure 2. In this case, low discount rates of the government, indicated by high levels of δ, in combination with a high probability ρ of a public rebellion can induce even a bad government to choose $e_1 = 0$, so that there is a pooling equilibrium.

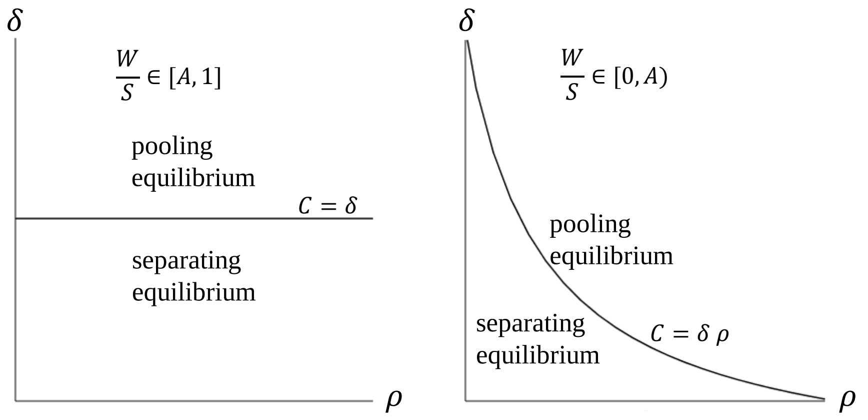

Figure 2: Dynamics of Insurrections

The reason is that high levels of δ imply a bad government to value period-two payoffs sufficiently high for keeping it interested in surviving in office at the end of period one. At the same time, a high ρ implies a high probability of the winning coalition to be freed from its loyalty trap by a public rebellion in the case of a choice $e_1 = e^*_{R_B}$ by a bad government. Should a public rebellion happen, then this turns dropping the government optimal for the winning coalition upon having first observed a choice $e_1 = e^*_{R_B}$ by the government and then a public rebellion. Hence, combinations of high values of δ and ρ are preconditions for a pooling equilibrium.[5]

By contrast, low probabilities of both δ and ρ induce a bad government to choose $e_1 = e^*_{R_B}$ while a good government will always choose $e_1 = 0$, which is common knowledge. Hence such a combination constitutes a separating equilibrium. There are two possible variants of a separating equilibrium that can be identified on the right-hand side of figure 2 and a third separating equilibrium on the left-hand side: Below the hyperbolic line on the right-hand side, a bad government either beliefs it can afford choosing $e_1 = e^*_{R_B}$ since the probability ρ of a public rebellion is low, or it heavily discounts period-two

[5]Note that, by the definition of C, a rise in the difference $V^R_{R_B} - V^R_{GW}$ shift the $C = \delta\rho$-line outwards, while a rise in $V^R_{R_B} - V_{LP}$ shifts the line inwards.

payoffs, so that δ is low, or both. Since a low probability of a public rebellion does not preclude a public rebellion, there are two possible variants: Should there not be a public rebellion resulting from the bad government's choice $e_1 = e^*_{R_B}$, then the winning coalition will remain stuck in its loyalty trap and abstain from ousting the government, even knowing that it is bad. We refer to this variant of the separating equilibrium as a "loyalty-trap equilibrium". By contrast, should there be indeed a public rebellion resulting from the bad government's choice $e_1 = e^*_{R_B}$, then this will free the wining coalition from its loyalty trap, so that ousting the bad government becomes its optimal choice. We refer to this variant of the separating equilibrium as an "accidental revolution".

The third possible variant of a separating equilibrium can materialize under the condition $\frac{W}{S} \in [A, 1]$, that is a regime structure that precludes the existence of a loyalty trap. Such a structure is depicted on the left-hand side of figure 2. Here, the probability of a public rebellion is not relevant for the winning coalition since it does not need such a rebellion for becoming freed from a loyalty trap. Under these conditions, a separating equilibrium is only feasible if a bad government discounts period-two payoffs as much as is possible for establishing $\delta > C$. In such a case, the government will be ousted with necessity since there is no loyalty trap that would deter the winning coalition from ousting the bad government. The reason why a bad government nevertheless chooses $e_1 = e^*_{R_B}$ is simply that it accepts being ousted since it weighs the maximum of period-one payoffs higher than any discounted value of possible period-two payoffs. We refer to this variant of a separating equilibrium as a "hazarded coup".

5 Empirical Implications

In our model, the winning coalition remains the single group that is capable of either keeping a government in office or ousting it. A mass revolt alone is not sufficient for a government to be overthrown, but its occasional outbreak may unintentionally prepare the ground for it by reducing the winning coalition's costs of doing so. The latter, in turn, may or may not take its chance, depending on whether its members expect a net increase in utility from a change in government or not. In a similar fashion, Hannah Arendt (2006, pp. 251-252) wrote in her famous book "On Revolution" in 1963:

> "The outbreak of most revolutions has surprised the revolutionist groups and parties no less than all others, and there exists hardly a revolution whose outbreak could be blamed upon their activities. [...] The part of the revolutionists usually consists not in making a revolution but in rising to power after it has broken out [...]."

The outbreak of a public rebellion, however, is only relevant under regimes that are characterized by a winning coalition that is relatively small compared to the selectorate.[6] In these cases, the contingency of a public rebellion has two potential effects: Firstly, it may deter even a bad government from a defective strategy and it may hence induce it to behave as if it were a good government. Secondly, it may enable a winning coalition to oust a government upon having updated its belief of the government's character by observing its defective action.

Note, however, that the contingency of a public rebellion can check the power of an autocratic government only to the extent that the probability ρ of a public rebellion is stable and known to all actors. This, however, is not very realistic. The eruption and, perhaps even more so, the strength of many if not most momentous public rebellions took even their active participants by surprise. When the East-German regime manipulated local elections in May 1989, they did not do that for the first time. To the contrary, this practice had prevailed over decades and few people had ever dared to openly object. This notwithstanding, the manipulations of May 1989 suddenly sparked an outburst of protests that neither the participants, nor the regime officials had expected. Head of the regime Honecker misjudged the situation even in

[6] Technically, this means $\frac{W}{S} \in [0, A)$.

two points: He did not expect the people to gather for the protests following the local elections and, further on, he did not expect his winning coalition to withdraw its loyalty in October 1989. This was not unique to the East-German revolution. Rather, many if not most spectacular revolutions fit into this pattern, including prominent cases like the Iranian Revolution of 1979, the Russian Revolution of 1917 or the French Revolution of 1789 (Kuran, 1989).

From a retrospective view these famous revolutions suggest the defective strategy of the respective incumbents to have been a disastrous mistake. Ex ante, however, this view may be misleading, particularly if we consider the many dictators that have successfully pursuing precisely such defective strategies without ever facing any serious protests, let alone regime changes. Within our framework, this would imply the probability ρ of a public rebellion to be rather low and possibly also unknown to the incumbent, his winning coalition, and the general public alike. The typical revolution would then be an "accidental revolution" as defined in table 4, but the underlying accidents of these revolutions are rare. They occur when incumbents are not aware that (this time) their defective strategy happens to spark a public rebellion although time and again it might hitherto have remained void of any serious consequence; and they occur because a discontented winning coalition takes the opportunity of a public rebellion for getting rid of a not (anymore) beloved incumbent without risking to be swept out of the winning coalition themselves.

Note that conjecturing the accidents behind revolutions to be rare is not in contradiction to the fact that there is indeed a number of empirical investigations finding public protests or unrest to raise the probability of a government to be overthrown by its winning coalition, at least in the case of military dictatorships (Casper and Tyson, 2014; Powell, 2012; Thyne, 2010). As far as this is the case, there is indeed an indirect relationship between public unrest and government overthrow. According to our approach, however, there are two filters between the degree of discontent of the public on the one hand and the power of an incumbent in a dictatorship on the other:

The first is the stochastic occurrence of mass revolts and the second is the interest of the winning coalition in ousting the incumbent, conditional on the former's observation of public unrest. Note, however, that only a particular and partly stochastic sequence of events may happen to pass these two filters and hence effectively endanger the incumbent's power. By the same token, only the expectation of such a particular combination of events may exert a disciplining effect on an autocrat's public policy. The latter of the two filters explains why security forces sometimes shoot at protesters and sometimes not. It also implies, however, that conceptions of revolution constraints that limit the scope for extractive strategies of autocrats by the threat of a public uprising are misleading at best, since the connection between exploitation and public uprisings is obscured by the two filters.

There are numerous historical examples of the interplay between the two filters, of which the difference between the security forces' reaction to the protests in Beijing and Berlin in 1989 is only one. Another prominent example is the difference in the reaction of the Russian soldiers to the uprisings in 1905 and February 1917. By the time of the 1905 revolution, the soldiers where still in a comparatively comfortable position within the Tsarist Regime. They hence decided to violently fight the rebels. Somewhat more than a decade later, they did not see a future for themselves in the perishing Russian Empire, particularly not after having been abused for a war that was both devastating and senseless in any respect from the outset. Consequently, when the masses gathered for the protests in February 1917, the Russian soldiers not only hesitated to shoot. Rather, one after the other eventually defected to the revolutionaries, and only that gave reason for Tsar Nicolas II to finally step down.

The revolts against the Assad regime in Syria that grew to a most violent and still ongoing civil war constitutes another example. This happened since the core of Assad's security forces remained loyal even in the light of increasingly violent fights and a growing degree of organization and armament of the different rebelling groups. For Assad's winning coalition, there was no way out of the loyalty trap.

Apart from that, our approach has some further empirical implications. Even if a defective policy of a regime happens to spark an accidental revolution, the group of the powerless will, if any, only be protected against the worst tax allocation there is in favor of an allocation that is best for the winning coalition but not best for the powerless. In the case of small winning coalitions, there will be not too big a difference between the latter two allocations, but there will be a big difference between the optimal tax allocation from the respective points of view of the winning coalition and the powerless. As a rule, the larger is the winning coalition, the smaller will be the difference between the winning coalition's and the powerless' optimal tax allocation, but note that large winning coalitions are a characteristic of democracies, not dictatorships. Hence the protective effect of a potential rebellion by the powerless is least effective in a dictatorship, but sadly this is the system in which such a protective effect would appear most desirable.

As shown in appendix A, the probability of a mass revolt is larger in regimes that are based on narrow winning coalitions. However, the effect they have for the powerless are lower there as compared to a regime based on a broader winning coalition. What is more, the winning coalition takes, if any, the chance of a rebellion for ousting an incumbent while safeguarding each member's position in the winning coalition. That, in turn, explains why rebellions face severe difficulties when it comes to changing the deeper roots of power in a country. More often than not, the same old elite appears behind the face of a seemingly revolutionized political power structure. This is another observation which our model can explain.

6 Conclusions

We have linked a threshold model to an agency model of the relation between an autocratic government and its supporting winning coalition in order to improve our understanding of the interplay between potential or manifest mass revolts, the conduct of winning coalitions and the power and conduct of an incumbent dictator. We have distinguished the unintended powerful side effect of expressive behavior of members of the general public from the deliberately exerted power of a winning coalition. The latter alone is the one that decides on ousting an incumbent in light of its observation of public unrest and in light of its updated belief on the character of the incumbent. As a result, there are two filters between the degree of relative deprivation of the general public on the one hand and an overthrow of a dictator on the other: the stochastic element in the outbreak of public unrest and the interest of the winning coalition in ousting the incumbent.

The model presented in this paper has a rich set of empirical implications. One is that it enables us to define the conditions under which public unrest can become threatening to a dictator in a way that either enhances the probability of an overthrow or imposes limits on the degree to which a dictator can exploit both the winning coalition and the general public. Another implication is that public unrest raises the probability of a winning coalition to decline loyalty to the incumbent dictator. This is compatible with the tentative observation in the literature of a higher frequency of military coups or other forms of enforced government changes in times of public unrest. The general public will nevertheless only weakly be protected by the threat of public unrest since deprivation is a necessary but not a sufficient condition for public unrest and since public unrest does only raise the probability of enforced government changes as far as an incumbent violates the interest of the winning coalition but not that of the general public.

Finally, the probability of public protests is higher but the effect of a change in government on the degree of relative deprivation of the general public is weaker in the case of narrow as compared to broad winning coalitions. Hence, public unrest resulting in ousted governments may be more frequent in regimes that are based on narrow winning coalitions, but the threat thereof is not likely to be functional with respect to an effective and efficient control of autocratic governments by the public.

References

Acemoglu, D. & Robinson, J. A. (2001). A theory of political transitions. *American Economic Review*, *91*(4), 938–963.

Acemoglu, D. & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Arendt, H. (2006). *On revolution*. Penguin.

Besley, T. (2007). *Principled agents?: the political economy of good government*. Oxford University Press.

Besley, T. & Kudamatsu, M. (2008). Making autocracy work. In *Institutions and economic performance*. Cambridge Mass.: Harvard University Press.

Bloch, P. C. (1986). The politico-economic behavior of authoritarian governments. *Public Choice*, *51*(2), 117–128.

Boix, C. (2003). *Democracy and redistribution*. Cambridge University Press.

Brennan, G. & Lomasky, L. (1993). *Democracy and decision: the pure theory of electoral politics*. Cambridge: Cambridge University Press.

Bueno de Mesquita, B. & Smith, A. (2009). Political survival and endogenous institutional change. *Comparative Political Studies*, *42*(2), 167–197.

Bueno de Mesquita, B. & Smith, A. (2010). Leader survival, revolutions, and the nature of government finance. *American Journal of Political Science*, *54*(4), 936–950.

Bueno de Mesquita, B., Smith, A., Siverson, R. M., & Morrow, J. D. (2005). *The logic of political survival*. The MIT Press.

Casper, B. A. & Tyson, S. A. (2014). Popular protest and elite coordination in a coup d'etat. *The Journal of Politics*, *76*(02), 548–564.

Gilli, M. & Li, Y. (2014). Accountability in one-party government: rethinking the success of chinese economic reform. *Journal of Institutional and Theoretical Economics (JITE)*, *170*(4), 616–645.

Gilli, M. & Li, Y. (2015). Coups, revolutions and efficient policies in autocracies. *European Journal of Political Economy*, *39*, 109–124.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 1420–1443.

Grossman, H. I. (1991). A general equilibrium model of insurrections. *American Economic Review*, *81*(4), 912–921.

Grossman, H. I. (1999). Kleptocracy and revolutions. *Oxford Economic Papers*, *51*(2), 267–283.

Gurr, T. R. (1970). *Why men rebel*. Princeton University Press.

Havel, V. (1985). The power of the powerless. *International Journal of Politics*, *15*(3/4), 23–96.

Hillman, A. L. (2010). Expressive behavior in economics and politics. *European Journal of Political Economy*, *26*(4), 403–418.

Kuran, T. (1989). Sparks and prairie fires: a theory of unanticipated political revolution. *Public Choice*, *61*(1), 41–74.

Kurrild-Klitgaard, P. (2004). The paradox of rebellion. In *The encyclopedia of public choice* (pp. 728–731). Springer.

Ledyard, J. O. (1984). The pure theory of large two-candidate elections. *Public choice*, *44*(1), 7–41.

Lichbach, M. I. (1998). *The rebel's dilemma*. University of Michigan Press.

Lohmann, S. (1993). A signaling model of informative and manipulative political action. *Amercan Political Science Review*, *87*(2), 319–333.

Lohmann, S. (1994). The dynamics of informational cascades: the monday demonstrations in leipzig, east germany, 1989–91. *World politics*, *47*(01), 42–101.

Marwell, G. & Oliver, P. (1993). *The critical mass in collective action*. Cambridge University Press.

McGuire, M. C. & Olson, M. (1996). The economics of autocracy and majority rule: the invisible hand and the use of force. *Journal of economic literature*, 72–96.

Myerson, R. B. (2008). The autocrat's credibility problem and foundations of the constitutional state. *American Political Science Review*, *102*(01), 125–139.

Olson, M. (1993). Dictatorship, democracy, and development. *American Political Science Review*, *87*(03), 567–576.

Powell, J. (2012). Determinants of the attempting and outcome of coups d' état. *Journal of Conflict Resolution*, *56*(6), 1017–1040.

Schelling, T. C. (1978). Micromotives and macrobehavior. *New York: Norton&Company.*

Svolik, M. W. (2009). Power sharing and leadership dynamics in authoritarian regimes. *American Journal of Political Science*, *53*(2), 477–494.

Svolik, M. W. (2013). Contracting on violence the moral hazard in authoritarian repression and military intervention in politics. *Journal of Conflict Resolution*, *57*(5), 765–794.

Thyne, C. L. (2010). Supporter of stability or agent of agitation? The effect of us foreign policy on coups in latin america, 1960–99. *Journal of Peace Research.*

Tullock, G. (1971). The paradox of revolution. *Public Choice*, *11*(1), 89–99.

Yin, C.-C. (1998). Equilibria of collective action in different distributions of protest thresholds. *Public Choice*, *97*(4), 535–567.

A Size of the Winning Coalition and Probability of a Rebellion

In this appendix, we demonstrate that the probability ρ of a process toward a rebellion equilibrium drops as the size of the winning coalition rises. We start by solving equation 2 in the text for a:

$$a = (1 + z)\gamma. \tag{A.1}$$

We take the definition $\gamma := \frac{g_i}{g} - 1$ of the degree of deprivation, use the average value $\bar{g}$ for the individual value g_i and substitute the result $\gamma = \frac{\bar{g}}{g} - 1$ into A.1. We then use the average value $\bar{z}$ instead of z and get:

$$a = (1 + \bar{z})(\frac{\bar{g}}{g} - 1). \tag{A.2}$$

We now substitute the optimal levels g_{R_G} and g_{R_B} of the public-goods supply g as chosen by the respective government. Remember that g_{R_G} and g_{R_B} are the optimal levels for a good and a bad government, respectively, as given in table 1. This yields:

$$a = \frac{(1 + \bar{z})(1 + \alpha)\bar{g}}{\alpha(1 + W)} - (1 + \bar{z}). \tag{A.3}$$

$$a = \frac{(1 + \bar{z})(1 + \alpha + \beta)\bar{g}}{\alpha(1 + W)} - (1 + \bar{z}). \tag{A.4}$$

for the good and bad government, respectively. In both cases, we obviously have $a'(W) < 0$. Combining this with $\rho'(a) > 0$ from equation 8 gives:

$$\rho'(a)a'(W) < 0. \tag{A.5}$$

The larger (smaller) the winning coalition, the lower (higher) is the probability of a process toward a rebellion equilibrium to be launched ■

B Perfect Bayes Equilibria (PBE)

B.1 Pooling Equilibrium

We start by assuming group GW to infer the existence of a pooling equilibrium implying $e_1(R_G) = e_1(R_B) = 0$ from the observation $e_1 = 0$. Given the latter, GW applies Bayes' rule and finds:

$$Pr(R_G|e_1 = 0) = \frac{Pr(e_1 = 0|R_G)\pi}{Pr(e_1 = 0|R_G)\pi + Pr(e_1 = 0|R_B)(1-\pi)} = \pi. \tag{A.6}$$

By contrast, GW cannot infer anything from a counterfactual observation $e_1 = e^*_{R_k}$. Any feasible assumption on $Pr(R_G|e^*_{R_k}) \in [0, 1]$ is hence admissible within the concept of a PBE. However, a good government would, by the assumption of our model, under any conditions be worse off when departing from $e_t = 0$. Therefore, we have:

$$Pr(R_G|e_1 = e^*_{R_k}) = 0. \tag{A.7}$$

Group $GW's$ payoffs from it's possible responses $r_{GW} \in \{DG, SG\}$ to $e_1 = 0$ over both periods are:

$$V^{GW}(DG|e_1 = 0) = V^{GW}_{GW} + V_{PDG},^{7} \tag{A.8}$$

and

$$V^{GW}(SG|e_1 = 0) = V^{GW}_{GW} + \pi V^{GW}_{GW} + (1-\pi)V^{GW}_{R_B}. \tag{A.9}$$

It can easily be demonstrated that $V^{GW}(SG|e_1 = 0) > V^{GW}(DG|e_1 = 0)$ $\forall \pi \in [0, 1]$. Hence, the best response by GW to $e_1 = 0$ is:

$$r^*_{GW}(e_1 = 0) = SG. \tag{A.10}$$

[7] See table 3 for the notation.

Upon having observed SG, a good government chooses $e_2 = 0$ and a bad government chooses $e_2 = e^*_{R_B}$, which generates the following payoffs:

$$V^{R_G}(e_1 = 0|SG) = (1+\delta)V^{GW}_{GW}, \tag{A.11}$$

and

$$V^{R_B}(e_1 = 0|SG) = V^R_{GW} + \delta V^R_{R_B}. \tag{A.12}$$

Note that the payoffs *off the equilibrium path* depend on whether a decision ($e^*_{R_B}$) by the government is followed by a peace equilibrium (pe) or a rebellion equilibrium (re). Considering A.7, we find group $GW's$ expected payoffs off the equilibrium path as:

$$V^{GW}(SG|e_1 = e^*_{R_B}) = 2V^{GW}_{R_B}, \tag{A.13}$$

if it continues supporting the government, as well as

$$V^{GW}(DG|e_1 = e^*_{R_B}, re) = V^{GW}_{R_B} + V_{RDG}, \tag{A.14}$$

if it drops the government in a rebellion equilibrium, and

$$V^{GW}(DG|e_1 = e^*_{R_B}, pe) = V^{GW}_{R_B} + V_{PDG}, \tag{A.15}$$

if it drops the government in a peace equilibrium. Things are straightforward in a rebellion equilibrium. Since it is easily demonstrated that $V_{RDG} > V^{GW}_{R_B}$, we also have that $V^{GW}(DG|e_1 = e^*_{R_B}, re) > V^{GW}(SG|e_1 = e^*_{R_B})$. Things are not that clear cut in a peace equilibrium, however. Here, we find $V^{GW}(DG|e_1 = e^*_{R_B}, pe) \geqslant V^{GW}(SG|e_1 = e^*_{R_B})$ only for sufficiently high values of W/S, namely for:

$$\frac{W}{S} \geqslant \frac{V^{GW}_{R_B} - V_{LP}}{\pi V^{GW}_{GW} + (1-\pi)V^{GW}_{R_B} - V_{LP}} =: A. \tag{A.16}$$

Taken together, we have the following best response of group GW to an observation $e_1 = e^*_{R_B}$ off the equilibrium path:

$$r^*_{GW}(e_1 = e^*_{R_B}) = \begin{cases} DG\ if\ re; \\ DG\ if\ pe \wedge \frac{W}{S} \in [A, 1]; \\ SG\ if\ pe \wedge\ \frac{W}{S} \in [0, A). \end{cases} \tag{A.17}$$

Note that the government does not know whether there will be a peace equilibrium or a rebellion equilibrium resulting from a choice $e_1 = e^*_{R_B}$. Based on the best responses by GW as given by equation A.17, a bad government's expected payoffs off the equilibrium path are thus:

$$V^{R_B}(e_1 = e^*_{R_B}|r^*_{GW}) = \rho(V^R_{R_B} + \delta V_{LP}) + (1-\rho)(1+\delta)V^R_{R_B} \quad \forall\ \frac{W}{S} \in [0, A), \tag{A.18}$$

and

$$V^{R_B}(e_1 = e^*_{R_B}|r^*_{GW}) = V^R_{R_B} + \delta V_{LP} \quad \forall\ \frac{W}{S} \in [A, 1]. \tag{A.19}$$

Option $e_1 = 0$ is the best response to r^*_{GW} by a bad government if the latter's payoff on the equilibrium path as given by A.12 is strictly higher than the payoff off the equilibrium path as given by A.18 or A.19, respectively. In particular, $e_1 = 0$ is the best response if:

$$V^R_{GW} + \delta V^R_{R_B} \geqslant \rho(V^R_{R_B} + \delta V_{LP}) + (1-\rho)(1+\delta)V^R_{R_B} \quad for: \frac{W}{S} \in [0, A), \tag{A.20}$$

or

$$V^R_{GW} + \delta V^R_{R_B} \geqslant V^R_{R_B} + \delta V_{LP} \quad for: \frac{W}{S} \in [A, 1]. \tag{A.21}$$

Condition A.20 is satisfied for all $(V^R_{R_B} - V^R_{GW})/\rho(V^R_{R_B} - V_{LP}) \leqslant \delta\rho$, while condition A.21 is satisfied whenever $(V^R_{R_B} - V^R_{GW})/(V^R_{R_B} - V_{LP}) \leqslant \delta$. We can hence summarize a bad government's best response $r^*_{R_B}$ to the winning coalition's best response r^*_{GW} as follows:

$$r^*_{R_B}(r^*_{GW}) = \begin{cases} e_1 = 0 \; if \; C \leqslant \delta | \frac{W}{S} \in [A, 1]; \\ e_1 = 0 \; if \; C \leqslant \delta\rho | \frac{B}{\rho} | \frac{W}{S} \in [0, A); \\ e_1 = e^*_{R_B} \quad otherwise, \end{cases} \tag{A.22}$$

where $C := (V^R_{R_B} - V^R_{GW})/(V^R_{R_B} - V_{LP})$. Finally, based on the best responses by GW as given by equation A.17, a good government's expected payoff off the equilibrium path is:

$$V^{R_G}(e_1 = e^*_{R_B} | DG) = V^{GW}_{R_B} + \delta V_{LP}. \tag{A.23}$$

Comparing this payoff to $V^{R_G}(e_1 = 0 | SG) = (1 + \delta) V^{GW}_{GW}$ on the equilibrium path, we immediately find that $(1 + \delta) V^{GW}_{GW} > V^{GW}_{R_B} + \delta V_{LP}$ and hence that $V^{R_G}(e_1 = 0 | SG) > V^{R_G}(e_1 = e^*_{R_B} | DG)$ under any conditions, so that.

$$r^*_{R_G}(r^*_{GW}) = \{e_1 = 0\}. \tag{A.24}$$

Summing up, we find a pooling equilibrium as described by the following strategies s^*_{GW} and $s^*_{R_k}$ of the winning coalition and the government:

$$s^*_{GW} = \{SG\}, \tag{A.25}$$

$$s^*_{R_k} = \begin{cases} e_t = 0 \quad if \quad k = G; \\ e_1 = 0; e_2 = e^*_{R_B} \quad if \quad k = B \wedge C \leqslant \delta | \frac{W}{S} \in [A, 1]; \\ e_1 = 0; e_2 = e^*_{R_B} \quad if \quad k = B \wedge C \leqslant \delta\rho | \frac{W}{S} \in [0, A) \end{cases} \qquad \blacksquare \tag{A.26}$$

B.2 Separating Equilibria

We start by assuming group GW to infer a separating equilibrium with $e_1(R_G) = 0; e_1(R_B) = e^*_{R_B}$ from successively observing first $e = e^*_{R_B}$ and then either a peace or a rebellion equilibrium. Based on $GW's$ assumption and applying Bayes' rule, the group finds:

$$Pr(R_B|e_1 = e^*_{R_B}) = 1. \tag{A.27}$$

The payoffs of the responses $r_{GW} \in \{DG, SG\}$ by group GW on $e_1 = e^*_{R_B}$ in a rebellion equilibrium (re) are:

$$V^{GW}(DG|e_1 = e^*_{R_B}, re) = V^{GW}_{R_B} + V_{RDG}, \tag{A.28}$$

and

$$V^{GW}(SG|e_1 = e^*_{R_B}, re) = 2V^{GW}_{R_B}. \tag{A.29}$$

It can easily be demonstrated that $V^{GW}(DG|e_1 = e^*_{R_B}, re) > V^{GW}(SG|e_1 = e^*_{R_B}, re)$ $\forall\ \pi \in (0,1)$, so that we find DG to be the best responses by group GW to $e_1 = e^*_{R_B}$ in a rebellion equilibrium.

The payoffs of the responses $r_{GW} \in \{DG, SG\}$ by group GW on $e_1 = e^*_{R_B}$ in a peace equilibrium (pe) are:

$$V^{GW}(DG|e_1 = e^*_{R_B}, pe) = V^{GW}_{R_B} + V_{PDG}, \tag{A.30}$$

and

$$V^{GW}(SG|e_1 = e^*_{R_B}, pe) = 2V^{GW}_{R_B}. \tag{A.31}$$

As in A.16, we have $V^{GW}(DG|e_1 = e^*_{R_B}, pe) > V^{GW}(SG|e_1 = e^*_{R_B}, pe)$ $\forall\ \frac{W}{S} \in [A, 1]$, so that we find the following best responses of group GW to $e_1 = e^*_{R_B}$:

$$r^*_{GW}(e_1 = e^*_{R_B}) = \begin{cases} DG \quad if \quad re; \\ DG \quad if \quad pe \wedge \frac{W}{S} \in [A, 1]; \\ SG \quad if \quad pe \wedge \frac{W}{S} \in [0, A), \end{cases} . \tag{A.32}$$

The payoffs stemming from $e_1 = e^*_{R_B}$ conditional on the winning coalition's best response for a good government and a bad government, respectively, are as follows:

$$V^{R_G}(e_1 = e^*_{R_B}|r^*_{R_B}) = \begin{cases} \rho(V^{GW}_{R_B} + \delta V_{LP}) + (1-\rho)(V^{GW}_{R_B} + \delta V^{GW}_{GW}) \\ \quad \forall\ \frac{W}{S} \in [0, A); \\ V^{GW}_{R_B} + \delta V_{LP} \quad \forall\ \frac{W}{S} \in [A, 1], \end{cases} \tag{A.33}$$

$$V^{R_B}(e_1 = e^*_{R_B}|r^*_{R_B}) = \begin{cases} \rho(V^{R}_{R_B} + \delta V_{LP}) + (1-\rho)(1+\delta)V^{R}_{R_B} \\ \quad \forall\ \frac{W}{S} \in [0, A); \\ V^{R}_{R_B} + \delta V_{LP} \quad \forall\ \frac{W}{S} \in [A, 1]. \end{cases} \tag{A.34}$$

Should group GW have observed $e_1 = 0$ instead of $e_1 = e^*_{R_B}$, the combination of the assumed separating equilibrium with Bayes' rule would yield:

$$Pr(R_B|e_1 = 0) = 0, \tag{A.35}$$

and GW would expect the following payoffs, conditional on it's response $r_{GW} \in \{DG, SG\}$:

$$V^{GW}(e_1 = 0|DG) = V^{GW}_{GW} + V_{PDG}, \tag{A.36}$$

and

$$V^{GW}(e_1 = 0|SG) = 2V^{GW}_{GW}. \tag{A.37}$$

It immediately follows that $V^{GW}(e_1 = 0|SG) > V^{GW}(e_1 = 0|DG)$, so that the best response to $e_1 = 0$ by GW is:

$$r^*_{GW}(e_1 = 0) = \{SG\}. \tag{A.38}$$

The payoffs from $e_1 = 0$ and $r^*_{GW}(e_1 = 0)$ for a good and a bad government, respectively, are then:

$$V^{R_G}(e_1 = 0|r^*_{GW}) = (1+\delta)V^{GW}_{GW}, \tag{A.39}$$

and

$$V^{R_B}(e_1 = 0|r^*_{GW}) = V^R_{GW} + \delta V^R_{R_B}. \tag{A.40}$$

Option $e_1 = e^*_{R_B}$ is a best response to r^*_{GW} for a good government, if $V^{R_G}(e_1 = e^*_{R_B}) > V^{R_G}(e_1 = 0)$, as given by equations A.33 and A.39. Since the latter is not valid for any combination of $\frac{W}{S}$ and ρ, we have:

$$r^*_{R_G}(r^*_{GW}) = \{e_1 = 0\} \;\; \forall \; \frac{W}{S}, \rho \in [0, 1]. \tag{A.41}$$

Option $e_1 = e^*_{R_B}$ is a best response to r^*_{GW} for a bad government, if $V^{R_B}(e_1 = e^*_{R_B}|r^*_{GW}) > V^{R_B}(e_1 = 0|r^*_{GW})$, as given by equations A.34 and A.40. The latter applies if $(V^R_{R_B} - V^R_{GW})/(V^R_{R_B} - V_{LP}) > \delta$ in the case of $\frac{W}{S} \in [0, A)$ or if $(V^R_{R_B} - V^R_{GW})/(V^R_{R_B} - V_{LP}) > \delta\rho$ in the case of $\frac{W}{S} \in [A, 1]$. We can hence summarize a bad government's best response $r^*_{R_B}(r^*_{GW})$ to the winning coalition's best response r^*_{GW} as follows:

$$r^*_{R_B}(r^*_{GW}) = \begin{cases} e_1 = e^*_{R_B} \;\; if \;\, C > \delta | \frac{W}{S} \in [A, 1]; \\ e_1 = e^*_{R_B} \;\; if \;\, C > \delta\rho | \frac{W}{S} \in [0, A); \\ e_1 = 0 \;\;\; otherwise, \end{cases} \tag{A.42}$$

In the case of a rebellion equilibrium and in the case of $e_1 = e^*_{R_B}$ plus $C > \delta | \frac{W}{S} \in [A, 1]$ and a peace equilibrium, the bad government's choice induces group GW to choose DG, according to condition A.32. In the case of $e_1 = e^*_{R_B}$ plus $C > \delta\rho | \frac{W}{S} \in [0, A)$ and a peace equilibrium, by contrast, $GW's$ choice is SG, according to condition A.32, even though GW knows the government to be of type $k = B$. We refer to the latter as a loyalty-trap equilibrium.

Summing up, we find a set of separating equilibria as described by the following strategies s^*_{GW} and $s^*_{R_k}$ of the winning coalition and the government:

$$s^*_{GW} = \begin{cases} SG \quad if \quad e_1 = 0; \\ DG \quad if \quad e_1 = e^*_{R_B} \wedge re; \\ DG \quad if \quad e_1 = e^*_{R_B} \wedge pe \wedge \frac{W}{S} \in [A, 1]; \\ SG \quad if \quad e_1 = e^*_{R_B} \wedge pe \wedge \frac{W}{S} \in [0, A), \end{cases} \tag{A.43}$$

$$s^*_{R_k} = \begin{cases} e_t = 0 \quad if \quad k = G; \\ e_t = e^*_{R_B} \quad if \quad k = B \wedge C > \delta | \frac{W}{S} \in [A, 1]; \\ e_t = e^*_{R_B} \quad if \quad k = B \wedge C > \delta\rho | \frac{W}{S} \in [0, A) \end{cases} \tag{A.44}$$

■

C Game Tree

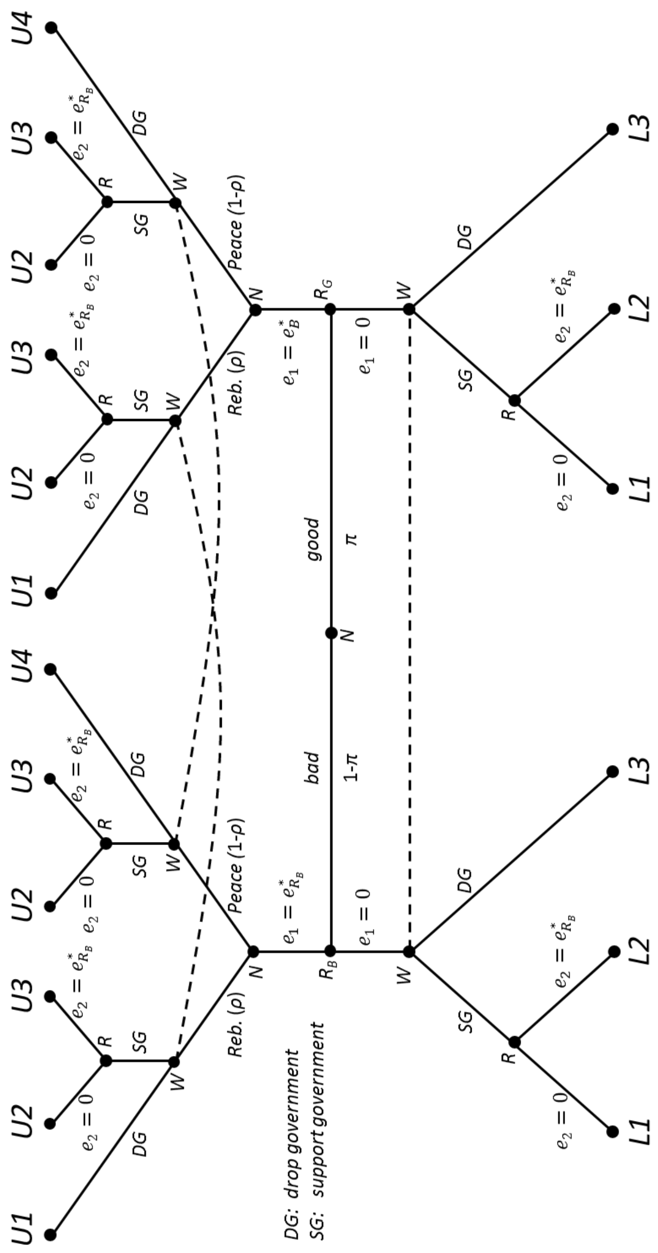

Figure A.1: Game Tree